

AIRT

AI-baserad Realtidsprognostisering av Trafikinformation

Daniel Jakobsson¹, Thomas Ludvigsson²,
Peter Lenaers³, Henrik Runeskog³, Katarina Elevant⁴,
Mattias Tiger⁵, Daniel Jönsson⁵

20 mars 2023

¹Trafikverket Evd (daniel.jakobsson@trafikverket.se)

²Trafikverket IKTiv (thomas.ludvigsson@trafikverket.se)

³Atea

⁴Fristående forskare, trafikväder

⁵Linköpings Universitet ({förnamn.efternamn}@liu.se)

Trafikverket FOI-projekt i portföljen *möjliggöra*

Sammanfattning

AIRT projektet etablerades för ta vidare den forskning från Projekt AI resultat och utveckla forskningsområden för att främja järnvägssystemets trafikinformation. Detta med syfte att förbättra trafikinformation som man idag operativt levererar till slutkund, våra resenärer inom järnvägssystemet. Forskningen planerades för att fokusera på att skapa *prediktioner av tidsåtgång för trafikstörande händelser, prediktioner av ankomsttid/avgångstid för tåg till station, Vädrets påverkan på järnvägssystemet* och *Visualisering av data och 3D*. Forskningen har gjorts tillsammans med akademi (LIU), näringsliv (ATEA) och offentligförvaltning (Trafikverket). Projektet startade med en godkänd projekt specifikation 2020-05-01 och nådde slutet på utförande fasen 2022-12-31. Genom inkrementell forsknings modell med stöd av olika metoder har vi utvecklat resultaten som presenteras i denna rapport. De framtagna prediktiva AI modellerna demonstreras ge en förbättrad prognostiserat jämfört med de prognoser som görs idag. Prediktiva analyser hjälper här att avgöra vad som sannolikt händer i en framtid. Det visar potentialen att i framtida samarbete med de operativa rollerna avsevärt förbättra kvaliteten på trafikinformation inom järnväg. För Trafikstörandehändelser har den iterativa modell utvecklingen resulterat till i snitt ca 7 minuter (11%) bättre än manuella prognoser fig 3:15 som representerar den manuella förmågan till prognos. Det finns potential till långsiktig förbättring och största steg framåt kan förmodligen göras genom att förbättra kvalitén av inputdata som kommer från den operativa processen, exempelvis skapa, klassificera och stänga en trafikstörande händelse. Ett annat problem är att det finns inte mycket data för flera av orsakskoderna. Tågs ankomst till station bygger på maskingenererat data vilket är stabilare då det kommer från järnvägssystemets signalsystem för tåg passager och den fastlagda tidtabellen. Det visar också att resultaten är mycket bättre i den prediktiva analysen. Forskningsområdet Vädrets påverkan i Järnvägs systemet har identifierat att antalet händelser öka vid definierat dåligt väder som innefattar en mängd väder parametrars kombinationer. Inom fokusområdet Visualisering av Data och 3D har en rapport skrivits och ett samarbete skapade vi en 3D värld med fokus på tågstäcka Stockholm - Göteborg med visualisering av prognoser från Trafikstörande Händelser och tågs ankomst och avgång från station. Det har varit fantastisk att i detta forskning projektet sett resultat växa fram som kan komma att betyda en hel verksamhetsförändring som kommer både våra operativa delar och resenärerna till nytta för en utvecklad trafikinformations tjänst.

Innehåll

1	Introduktion	3
1.1	Bakgrund	3
1.2	Struktur och syfte	4
1.3	Fokusområden	4
2	Bakgrund	6
2.1	Maskininläring	6
2.1.1	Maskininlärningsmetoder	6
2.1.2	Parameteranpassning	7
2.2	Data readiness	7
2.2.1	C-bandet	7
2.2.2	B-bandet	8
2.2.3	A-bandet	8
2.3	Relaterade arbeten	8
3	Trafikpåverkande händelser	10
3.1	Introduktion	10
3.2	LUPP data	10
3.3	Väderdata	10
3.3.1	VViS-stationer	10
3.3.2	Interpolation av väderdata	11
3.4	Datautforskning	11
3.4.1	Kategoriska variabler	13
3.4.2	Kontinuerliga variabler	14
3.4.3	Interntext	16
3.5	Modellering	18
3.6	Resultat	19
3.7	Diskussion och slutsatser	20
4	Ankomsttid för tåg	22
4.1	Introduktion	22
4.2	Datautforskning	22
4.3	Modellering	23
4.3.1	Features och Hyperparameterisering	23
4.3.2	Feature Importance och modellförklaring	26
4.3.3	Framtida prognoser	28
4.4	Resultat	29
4.4.1	För Nästkommande Station	30
4.4.2	Framtida prognostering	32

5 Vädrets påverkan på järnvägssystemet	40
5.1 Väderdata	40
5.1.1 Geospatial anpassning och kartläggning av delsträckor i järnvägsnätet	41
5.2 Metoder och vetenskaplig metodik	42
5.2.1 Metodologiska utmaningar och vetenskaplig metodik	42
5.2.2 Kvalitativa och kvantitativa metoder	43
5.2.3 Analysverktyg	44
5.3 Variabler associerade med händelser	44
5.3.1 Merförsening	45
5.3.2 Trafikpåverkans längd	45
5.3.3 Händelsefrekvens	45
5.3.4 Text	45
5.4 Meteorologiska variabler	45
5.4.1 Järnvägsväder	45
5.5 Resultat	47
5.5.1 Dataunderlaget	47
5.5.2 Vädertyper	47
5.5.3 Samband mellan väder och händelser	49
6 Visualisering av trafiken och händelser i 3D	57
7 Diskussion	61
7.1 Förslag på hur prediktioner kan integreras i verktyg	61
8 Slutsatser och rekommendationer för framtida arbete	62
8.1 Trafikhändelser	62
8.2 Tågs ankomst till station	62
8.3 Vädrets påverkan på järnvägssystemet	64
8.3.1 Framtida "Järnvägsväder" och vädertyper	64
8.3.2 Kartläggning av järnvägsnätet	64
8.3.3 Introduktion av nya variabler	64
8.3.4 Geografiska skillnader	64
8.4 Framtidens visualisering genom digitalatvillingar	65

Kapitel 1

Introduktion

Trafikverket har många viktiga samhällsfunktioner och ansvarar för samhällsviktig infrastruktur så som järnvägsnätet. Trafikverket har mer specifikt uppgifterna att underhålla järnvägsnätet, koordinera och informera tågoperatörer i deras användning av järnvägsnätet samt att ge relevant trafikinformation till resenärer som reser med dessa tågoperatörer. Kännedom om hur lång tid en trafikstörande händelse kan tänkas ta innan trafik fortlöper som vanligt är därför högst relevant. Detsamma gällande när tåg förväntas ankomma och avgå från stationer. Prognoser av dessa tider görs av trafikinformatorer, som endast har basala automatiska förslag och stöd så som medelvärdet för tidigare år. Genom bättre prediktiva modeller kan trafikinformatorer stärkas så att de ges förutsättningar att göra bättre prognoser. Detta kan i förlängningen leda till snabbare, exaktare och mer relevant trafikinformation till resenärer, tågoperatörer, underhållsarbetare och trafikverkets interna processer i stort. Den potentiella samhällsnyttan detta kan innebära är således mycket stor.

I projektet *AI-baserad Realtidsprognostisering av Trafikinformation* (AIRT) har forskare, specialister och domänexperter fördjupat sig kring framtagandet av prediktiva modeller för att förutsäga *tidsåtgång för trafikstörande händelse* och *ankomsttid/avgångstid för tåg till station*. Projektet har även forskat på *vädrets påverkan i järnvägssystemet*. Det inkluderar ett intensivt arbete för att lyfta Trafikverkets *data readiness* inom en avgränsning lämplig för projektet genom att använda *visualiserings tekniker* som ökad förståelse av data. Detta för att ta fram lämpliga tekniker och metoder från Artificiell Intelligens (AI), specifikt från Maskininlärning (ML), kunnat utvecklas och utvärderas för dessa prognostiseringsproblem. De framtagna prediktiva modellerna demonstreras ge en förbättrad prognostisering jämfört med de prognoser som idag görs. Genom förankring i verksamheten och dess domänexperter och utövare har det troliggjorts att de förväntas ge ett förbättrat stöd för trafikinformatorer att göra bättre prognoser. I slutet av projektet togs ett konkret initiativ där vi tog vara på analys resultaten i form av prognoser från tidsåtgång för trafikstörande händelse och ankomsttid/avgångstid för tåg till station och presenterade det i en 3D miljö.

AIRT är ett FOI-projekt finansierat av forskningsportföljen *möjliggöra* inom Trafikverket. Projektet är inriktat på målområde *Mer användbar trafikinformation och säkrare prognoser*.

1.1 Bakgrund

AIRT är baserad på det tidigare forskningsprojektet *Projekt AI*[1], som drevs av Trafikverket i samarbete med Linköpings Universitet. Projektet lade grunden för arbetet inom prognos för tidsåtgång för trafikstörande händelser.

Projekt AI tog fram ett kunskapsunderlag för framtida realtidsprognoser av tågtrafik med avseende på enskilda tågindividers ankomsttid respektive enskilda trafikstörande händelsers varaktighet. Trafikverket har stora mängder historisk data över trafik och trafikhändelser. Målet med projektet var att undersöka AI-tekniker och framför allt maskininlärnings-tekniker som utifrån denna data ska ha möjlighet att förbättra prediktioner avseende ankomsttid och störningslängd jämfört med prediktioner så som de görs i verksamheten idag.

Arbetet inom Projekt AI ägnades till största del om att lyfta lämpliga datamängder till hög data readiness-nivå och mynnade ut i rekommendationer kring fortsatta AI-projekt och fortsatt arbete med data readiness inom Trafikverket. En första ansats med maskininlärningsmodeller utvärderades även för prediktion av tidsåtgång för händelse. Projektet la grunden för AIRT fortsatta forskning med utökade forsknings områden som , beskrivs i kapitel Struktur och syfte.

1.2 Struktur och syfte

AIRT står för AI-baserad Realtidsprognostisering av Trafikinformation. Den första delen kommer från det övergripande begreppet för maskinlärning, modeller med förmågan att ge förutsägbara/skattande prognoser. Realtid relaterar till den förväntade framtids beteendet för prognostiseringen där tåg och händelser respektive rörpåsig och sker i realtid. Trafikinformation är här relaterat till Järnvägssystemet.

Projektet startade med en godkänd projektspecifikation 2020-05-01 och nådde slutet på Genomförande Realiserings fasen, TG4 2022-12-31. Överlämningsfasen och avslutning fram till TG5 utfördes till 2023-03-31. Tollgate, TG begreppen förklaras mer i XLPM projektmodell.

AIRT är uppdelat i fyra fokus områden som en struktur för projektet. Samarbete och resurs delning mellan fokus områdena har varit flexibelt men personer har fått tagit ledarskap inom ett specifikt fokusområde.

Dessa fokusområde är:

- Prediktion av tidsåtgång för trafikstörande händelse
- Prediktion av ankomsttid av tåg till station
- Vädrets påverkan på järnvägssystemet
- Visualisering av trafiken och händelser i 3D

Projekt AIRT är ett projekt drivet av Trafikverket i samarbete med Atea och Linköpings Universitet. Projekt AIRT avser att ta fram ett kunskapsunderlag för framtida realtids-prognoser av tågtrafik med avseende på *enskilda tågindividers ankomsttid* respektive *enskilda trafikstörande händelsers varaktighet*. Trafikverket har stora mängder historisk data över trafik och trafikhändelser. Forskning kring vädrets påverkan på järnvägssystemet för att identifiera på verkans faktorer inom väderdata. Visualisering har användas i projektet för att se på data sätt men också för att visualiserare resultat och användning av 3D teknik. Målet med projektet är att undersöka AI-tekniker och framför allt maskininlärningstekniker som utifrån denna data ska ha möjlighet att förbättra prediktioner avseende ankomsttid och störningslängd jämfört med prediktioner så som de görs i verksamheten idag. I detta projekt deltar ett flertal personer från Trafikverket, Atea och Linköpings Universitet. Från Trafikverket deltar Daniel Jakobsson, Thomas Ludvigsson, Johan Klingberg, Kristoffer Rönnerberg och Tomas Rendalen. Från Atea deltar Peter Laeners och Henrik Runeskog. Från Linköpings Universitet deltar Mattias Tiger från AIICS/ReaL (ledd av Fredrik Heintz) och Daniel Jönsson från MIT (ledd av Anders Ynnerman). Katarina Elevant deltar som fristående forskare.

1.3 Fokusområden

Prediktion av tidsåtgång för trafikstörande händelse. En stor mängd av distinkta trafikstörande händelser kan inträffa och som leder till trafikstopp eller trafikförseningar. Störningsområdet kan sträcka sig förbi flera stationer. Störningen kan orsaka ett lågfartsområden eller helt stopp på ett eller flera spår. En stor mängd faktorer förväntas vara med att påverka tidslängden av olika typer av störningen. Bland dessa så är väder troligt att vara en stark faktor, men även när (timme på dygnet, dag i veckan, dag på året) och var (nära städer, nära vägar, i norr eller i södra Sverige) har troligtvis högt informationsvärde. Det finns även rapporter i form av ostrukturerad text att tillgå.

Prediktion av ankomsttid av tåg till station. Då tågnätet är begränsat med enkelspår eller dubbelspår och nätet är vältrafikerat så påverkas ankomsttiden för ett givet tåg potentiellt sett av många andra tåg. Problemet är därför styrt av den struktur som finns i tågnätet, hur stationer är ihopkopplade och hur förbindelserna ser ut mellan respektive station. Bra prediktioner förväntas kräva att andra tåg tas i beaktande och indirekt deras i sin tur förväntade ankomsttider utmed hela tågnätet. Slutligen så påverkar trafikstörande händelsers tidslängd ankomsttiden mellan stationer.

Vädrets påverkan på järnvägssystemet Som ett delmål i projektet undersöktes samband mellan väder och händelser i järnvägssystemet, samtidigt som arbetet inom detta fokusområde utgjorde stöd för den övriga forskningen inom AIRT. Samband mellan väder och händelser i järnvägsnätet har tidigare inte utretts i någon större omfattning även om man inom Trafikverket dagligen arbetar med väderprognoser som beslutsunderlag. Fokusområde Vädrets påverkan på järnvägssystemet studerade samband mellan å ena sidan kombinationer av meteorologiska

variabler och variabler som beskriver konsekvenserna av vädret på marken och å andra sidan variabler associerade med händelser i järnvägssystemet.

Samband mellan de adresserade problemen. Prediktion av tidsåtgång för trafikstörande händelse och Prediktion av ankomsttid av tåg till station hänger ihop på så vis att ankomsttid kan påverkas av förekomsten av trafikstörande händelse. Det är dock inte nödvändigtvis så att dessa två problem bäst löses i ordning. Inte heller är det nödvändigtvis så att prediktionen om tidsåtgång för händelse är en lämplig input till uträkningen av prediktion av ankomsttid, då händelsens plats inte är precis och givet de olika förutsättningar som finns för att åka in tid fram till närmaste station. Vädrets påverkan av järnvägssystemet utgör i sin tur input till respektive prediktionsproblem. Det är däremot inte nödvändigtvis så att just dessa prediktionsproblem stärks i betydande grad av de tillgängliga väderinput inom ramen för projektet.

Det är också möjligt att prediktion av ankomsttid går att meningsfullt förbättra även utan att hänsyn tas till trafikstörande händelser. I fallet då ett tåg är försenat (oavsett anledning) så kan det fortfarande gå att dra slutsatser om hur denna försening påverkar övrig tågtrafik och därmed ge bättre prediktioner för ankomsttid för alla tåg i nätet. Tas inte trafikstörande händelser med så agerar dessa som en bruska (skapar högre osäkerhet) kring de predikterade ankomsttiderna. Hur stor denna tillförda osäkerhet blir går inte att säga innan utförlig modellbyggnad och dataanalys har utförts.

Att vädret har en påverkan på järnvägsnätet är känt, dock finns andra orsakssamband och variabler oberoende av väder som kan påverka utfallet. Det betyder att vädervariablerna utgör endast en del av problemet. Vidare kan väderrelaterade störningar i sin natur avvika från medelvärden och sambanden på så sätt vara svårare att studera med AI metoder som utgångspunkt.

Visualisering Tekniker för att undersöka, se relationer, presentera stora datamängder krävs visualisering så vi kan uppfatta tolka information och dra slutsatser. Visualisering av indata och resultat i analysarbetet har varit en aktivitet i alla fokusområden. Inom *fokusområdet visualisering* undersöks de barriärer som det finns utmaning kring utveckling av avancerad visualisering [2]. Att kunna se och följa tågs rörelser med den information som relaterar till järnvägssystemet tex. trafikstörande händelser och prognos på tågen för att ge ökar förståelsen genom den informationsökning som ges. Vi har fokuserat på hur förståelsen för analysresultaten inom trafikstörande händelse och tåg i tid kan visualisera i 3D miljö som ett exempel på hur information kan upplevas.

- Bakgrund
- Vilka som är med
- Hur vi har arbetat i förhållande till tidigare AI-projekt på Trafikverket

Kapitel 2

Bakgrund

2.1 Maskininlärning

Vissa funktionalitet kan vara oerhört svår att specificera för en dator hur den ska utföras. Samtidigt kan det finnas, eller enkelt göra tillgängligt, mängder med exempel data på hur funktionen ska fungera. Kanske rent av så mycket data att det är ohanterligt för en människa att gå igenom och hitta önskade samband. Att genom domänkunskap om problem och tilltänkt användning anpassa en matematisk modell till att fånga upp (lära sig) funktionens natur (relationen mellan input och output) från data är *maskininlärning*. Konkret så handlar det om att hitta lämpliga värden för de parametrar som är fria i den matematiska modellen, där olika värden ger olika relation mellan input och output.

För att få ett bra resultat genom maskininlärning krävs bra datakvalité och datakänedom (*data readiness*, Kapitel 2.2). Detta då det dels informerar valet av modell och träningsmetodik, men också då maskininlärning inte kan lära något från en datamängd som inte redan finns som (gömd) information i datamängden i fråga. Om inlärningen gått som tänkt så fångar modellen hela den tilltänkta funktionaliteten, även för sådana input som inte var en del av (inte observerade / insamlade) i datamängden som modellen tränades på. Detta kallas generalisering. Målet är att kunna göra förutsägelser (så kallade prediktioner) som är tillräckligt tillförlitliga för att vara användbara.

Maskininlärning är en del av området Artificiell Intelligens (AI). För att få maskininlärning, framför allt den inlärda funktionaliteten, att bli faktiskt användbar i skarp verksamhet så krävs typiskt en kombination av många olika AI-tekniker utöver maskininlärning. Exempel på sådana tekniker är kunskapsrepresentation, slutsatsdragning (exempelvis för att övervaka att den inlärda modellen fortsätter fungera som tänkt), planering, beslutsfattande och inte minst interaktion med de mänskliga partners vars förmågor AI-systemet är tänkt att förstärka. Grundlig användning av statistik för att sammanfatta viktiga sammanhang och egenskaper av data, inlärningsresultat och prediktionsutfall är nödvändigt för att skapa förståelse för AI-systemets förmågor. Visualisering är fundamentalt för att kunna åskådliggöra och kommunicera detaljer likväl som stora trender är essentiellt för en god interaktion mellan operatör och AI-system. Rätt val av lämpliga AI och maskininlärningstekniker, så väl som en gedigen integration i ett verksamhets-sammanhang, är avgörande för att få praktiska resultat av värde för en skarp verksamhet.

2.1.1 Maskininlärningsmetoder

Maskininlärning har upplevt en stort boom senaste årtiondet, mycket tack vare stora framsteg inom *deep learning* genom artificiella neuronnät (NN) med många lager (så kallade djupa) som möjliggjort landvinningar av inlärning över ostrukturerad data. Ostrukturerade och semi-strukturerade datamängdstyper inkluderar ljud [3, 4], bilder [5, 6, 7] och text [8, 9, 10]. Sådan semi-strukturerad data har varit frustrerande svår att få användbar med tidigare maskininlärningstekniker, i huvudsak på grund av svårigheten att ta fram bra så kallade *features*. En feature är vanligtvis en tillrättalagd datatyp som fångar något informativt hos en datapunkt (ett exempel i från en datamängd). Exempelvis färgen på ett tåg istället för ett gäng pixlar i en bild, eller strömstyrkan genom en kontaktor i stället för ljudet av ljusbågen, eller huruvida ett ord är positivt eller negativt laddat istället för en handfull bokstäver utan ytterligare upplysning.

En grundläggande anledning till att *deep learning* varit så fruktsam är att inlärningen består i att både lära en

bra feature-representation och att göra bra prediktioner - samtidigt. Ett annat stort framsteg inom *deep learning* är att det blivit möjligt att ta fram mycket stora modeller, tränade på enorma datamängder. Genom att ta de första lagren av dessa NN så får man väldigt kraftfulla features (så kallade *deep features* vilket möjliggör användning av andra kraftfulla maskininläringstekniker än NN för data som tidigare var obrukbar.

Många organisationer har i första hand tillgång till stora mängder strukturerad data. Det är data som upp-kommer som data-spår från produktionssystem (så som *signal*), eller som matas in av operatörer (så som tågplan eller manuell prognos). Strukturerad data har typiskt redan starka och informativa features, så ytterligare feature-representation är därför onödig. Samtidigt som de välkända framstegen gjorts inom *deep learning* så har även stora framsteg gjorts inom maskininläringstekniker för strukturerad data. XGBoost [11] är ett exempel på en så-dan teknik som har uppnått vinnande prestanda på en mycket stor mängd av olika maskininlärningsutmaning-ar [11] av relevans för såväl företag som myndigheter. XGBoost, tillsammans med de likvärdigt högpresterande LightGBM [12] och CatBoost [13] är *gradient boosting*-metoder [14]. Dessa tre industriella metoder tävlar inbördes om att vara de mest användbara för prediktionsuppgifter över strukturerad data idag, då de ofta slår de mycket långsammare och svårhanterliga artificiella neuronäten när det handlar om tabulär data [15, 16, 17].

2.1.2 Parameteranpassning

Tidigare har modellparametrar introducerats. Det vill säga parametrar vars värde hittas genom anpassning (inlä-rning) av maskininlärningsmodellen till data. Utöver dessa finns också inlärningsparametrar (så kallade *hyperpa-rametrar*) som kontrollerar hur själva inläringen går till. Att ha bra värden för inlärningsparametrar är viktigt för att modellen ska kunna läras bra och därmed för att kunna bli användbar. Vad som är bra inlärningsparametrar beror på egenskaper hos det sammanhang som modellen ska användas i, något som förhoppningsvis finns väl re-presenterad i den tillgängliga datamängden som används för inläring. Utan en sådan anpassning är det svårt att få ut bra prestanda ur maskininläring.

En stor utmaning är att det inte går att söka efter bra värden för modellparametrar och inlärningsparametrar samtidigt, då inlärningsparametrarna påverkar vad som är bra modellparametrar. Därför krävs det att varje under-sökt uppsättning av inlärningsparametrar utvärderas genom en hel inläring av samtliga modellparametrar. Därav är parameteranpassning av inlärningsparametrar oerhört resurskrävande. Det är därför viktigt att parameteran-passningen sker smart och genom att testa så få möjliga parametervärden som möjligt - så kallat sampeleffektiv.

Bayesiansk Optimering (BO) [18] är en oerhört sampeleffektiv parameteranpassningsmetod som utnyttjar ro-bust maskininläring för att effektivt leta efter de bästa parametervärdena med så få experiment som möjligt. HyperOpt [19] är ett av flera välanvända mjukvaror som implementerar BO. Bayesiansk optimering används till största del inom maskininläring, då det är mycket effektivare än rutnätssökning eller slumpsökning, men det har också hittat många användningsområden utanför maskininlärningsfältet [18, 20]

2.2 Data readiness

Att nå hög *data readiness* är en av flera viktiga steg på vägen mot att kunna sjösätta AI och maskininläring i skarp verksamhet [21]. Olika former av data readiness har studerats över åren för att möta behoven inom Big Data [22], för data att senare användas för maskininläring generellt [23] och för AI- och maskininlärningsprojekt i synnerhet [24].

Neil Lawrence definierar tre band C,B,A [24] som beskriver hur redo ett visst dataset är för AI och Machine Learning, så kallat *Data Readiness* på engelska. C-bandet är det första och minst förberedda bandet och C1 är det mest förberedda tillståndet hos C-bandet. De olika tillstånden följer på varandra, C*-C1-B*-B1-A1, där A1 gäller data som är som mest förberedd för meningsfull användning. C* och B* här indikerar att det kan finnas flera till-stånd för C-bandet respektive för B-bandet, men att ett förslag på antal inte ges av Neil Lawrence [24]. Vi behöver nå motsvarande A1 för att i detalj kunna föreslå meningsfulla approacher och algoritmer för de två problemen.

2.2.1 C-bandet

C-bandet avser tillgängligheten till data. Det inkluderar ett flertal grundläggande aspekter som berör vilken data som finns, hur den är sparad, var den är sparad liksom hur den får användas. De huvudsakliga aspekterna kan delas upp i följande fyra kategorier (utan inbördes ordning):

- C₁ Huruvida påstått existerande data faktiskt finns sparad.
- C₂ Vad för format data finns sparad som. Format avser här exempelvis filtyp.
- C₃ Vilka juridiska begränsningar som finns för att få tillgång till och använda sig av data.
- C₄ Begränsningar på datatillgång så som hur det lagras, hur det får hämtas, hur eller var det får bearbetas etc.

Uppfylls inte C₁-C₄ för en viss datamängd så är det inte möjligt att gå vidare till B-bandet då data inte går att bearbeta eller undersöka på ett meningsfullt vis.

2.2.2 B-bandet

B-bandet avser aspekter gällande informationsinnehållet i data. Är data korrekt, reflekterar det den verklighet vi vill att det representerar och hur är osäkerheter i data hanterade? De huvudsakliga aspekterna kan delas in i 6 kategorier (utan inbördes ordning):

- B₁ Vilka enheter enskilda attribut har. Exempelvis sekunder eller timmar, absolut eller relativ tid.
- B₂ Om data har aggregerats eller förbehandlats och hur detta har skett. Viktig information kan ha tappats bort och slutsatser som dras kan vara påverkade eller inkorrekta.
- B₃ Huruvida det saknas data på vissa ställen i datamängden och hur det hanteras. Exempelvis att det förekommer konstgjorda värden (e.g. NaN, tom sträng eller negativt tal) som markerar detta.
- B₄ Huruvida det finns felaktig data och om det är markerat på något vis.
- B₅ Huruvida det finns osäker data och hur osäkerheten karaktäriseras.
- B₆ Huruvida det finns trender i dessa egenskaper hos data, exempelvis att data slutat vara (lika) felaktig eller brusig efter ett visst datum.

Datautforskning (*Exploratory Data Analysis*) används vanligtvis för att utforska dessa aspekter. Som en del av detta kan så väl generella statistiska metoder som specifika maskininlärningsmetoder användas för att studera olika frågeställningar och egenskaper.

Målet är att bygga upp en generell idé om begränsningarna i data och om vad för information som förväntas finnas i detta data. Det ska bli tydligt hur tillförlitlig och korrekt informationen är och vad som verkar vara möjligt att göra med den.

2.2.3 A-bandet

Först i A-bandet är en specifik användning i fokus, med frågeställningen om det givna data och dess informationsinnehåll är tillräckligt för att nå ett visst resultat. I detta projektet handlar det främst om trafikpåverkanslängden för händelser samt ankomsttid för tåg längs en sträcka.

2.3 Relaterade arbeten

Det finns inga publicerade arbeten till författarnas kännedom där de aktuella problemställningarna har inte studerats på den skala och givet den typ av datamängd förut till författarnas kännedom. Vissa relaterade problemställningar eller sammanhang har studerats och viss begränsad inspiration kan tas från dessa arbeten i fortsatta projekt då applikation av AI-tekniker har ett större fokus än det förarbete av datautforskning och lyft av datakvalité som blev huvudfokus för detta projekt.

Prediktion av passagerarefterfrågan och passagerarflöde för intercitytåg [25][26] har gjorts med hjälp av statistiska metoder så som ARIMA och SARIMA. Sådana prediktioner har även studerats för pendeltåg [27].

Mer relaterat till projektets aktuella problemställningar är prediktion av tågtrafik, som har studerats i Tokyo [28] där författarna använt djupa neurala nätverk baserade på CNN- och LSTM-arkitekturer. Det finns även arbeten på

prediktion av ankomsttid och avgångstid för tåg i Nederländerna [29], där senare arbeten även studerat prediktion av väntetider och tidslängden på tågfärder över ett intervall av 3 månader [30].

I biltrafik har studier från USA visat att mer än 25% av trängsel i trafiken orsakas av trafikstörande engångshändelser [31]. Författarna använder sig av Bayesiansk maskininlärning och neurala nätverk för att prediktera tiden det tar att åtgärda trafikhändelser innan normalt flöde kan återtas.

Kapitel 3

Trafikpåverkande händelser

3.1 Introduktion

En händelse som inträffar kan påverka trafiken på en viss sträcka. Allt från en hastighetsnedsättning till och med att trafiken är helt inställt. I denna del undersöker vi om vi kan bättre prediktera hur länge trafiken kommer att vara påverkat när en trafikstörande händelse inträffar.

Det finns många olika faktorer som påverkar längden av trafikpåverkande händelser. Faktorer som beskriver själva händelsen som vilken typ av händelse och var i Sverige händelsen inträffar är kanske självklart, men här undersöker vi också andra faktorer som till exempel väderomständigheter.

3.2 LUPP data

Data från historiska händelser kommer från LUPP databasen. Databasen innehåller information om varje händelse och varje gång en händelse får en uppdatering, ett nytt rad läggs till med information om händelsen. Efter datarensning finns 36296 händelser kvar från 2015 – 02 – 03 till och med 2021 – 05 – 18.

3.3 Väderdata

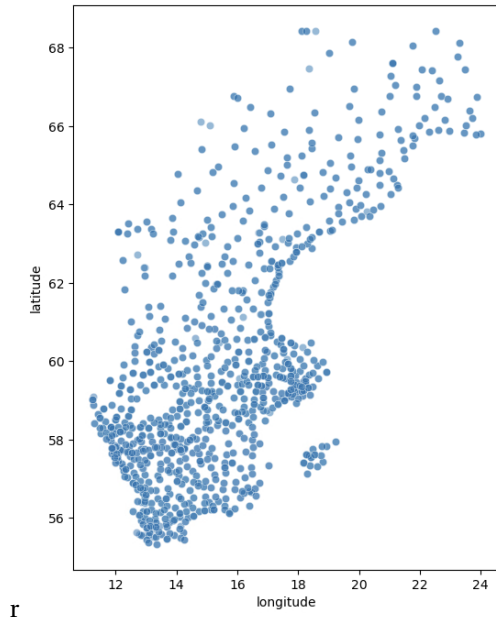
Vädret är en viktig faktor för trafiken på vägar och järnvägar. Säkerhet och framkomlighet är beroende av vår förmåga att kunna förutsäga, förebygga och varna för påverkande temperatur och nederbörd på vägar och järnvägar. Halka, särskilt frosthalka, osynlig ishalka på vägarna, och snöhinder på vintern. Kraftiga regn och åska på sommaren. För att få noggran och lokal information, har Trafikverket placerat ut egna väderstationer genom hela landet, ett så kallat Väg Väder Informations System med förkortningen VViS.

3.3.1 VViS-stationer

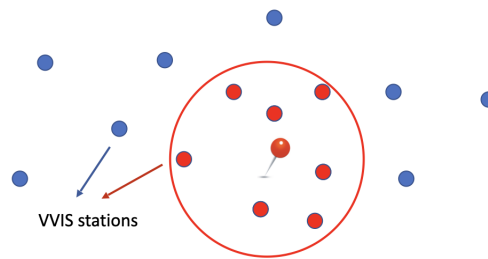
VViS-stationer mäter på regelbundna interval vädervariabler så som temperatur, vindriktning, vindstyrka och nederbörd. Stationerna finns fördelat över hela Sverige (Figur 3.1, men är lite glesare i Norra Sverige).

Data finns från 2013-09-02 till och med 2021-05-31 och innehåller information om 902 VViS-stationer. Felvärden filtreras bort så att data som behålls uppfyller följande kriterier för respektive vädervariabel:

- $-40 \leq \text{temperatur} \leq 50$
- $-35 \leq \text{daggpunkt} \leq 25$
- $0 \leq \text{vindstyrka} \leq 45$
- vindriktning $\neq -9$
- nederbördstyp $\neq \pm 9$



Figur 3.1: Lokalisering av VViS-stationer.



Figur 3.2: Interpolering av VViS data.

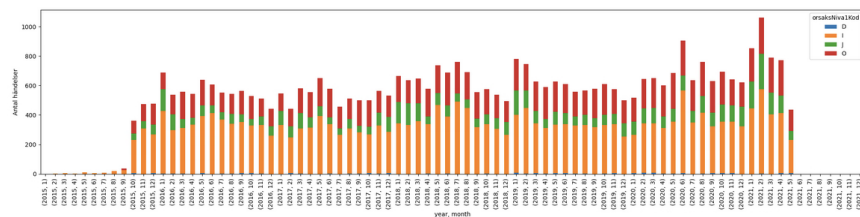
- nederbörds mängd ≥ 0
- luftfuktighet ≥ 0

3.3.2 Interpolation av väderdata

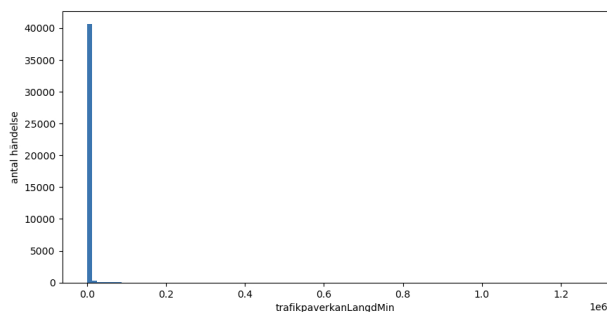
En händelse inträffar aldrig exakt där en VViS-station finns. Därför behöver vi beräkna lokala vädervariabler på drabbat plats. Det gör vi genom en interpolering. Det finns olika sätt hur man kan göra men vi beslutade att ta ett viktat snittvärde av de närmaste 7 VViS-stationerna med vikterna omvänt proportionell med distansen till platsen för händelsen, handelseplats. Så ju närmare stationen ligger, desto mer påverkar det beräknade värden. Metoden har en bra balans mellan noggrannhet och komplexitet. Så beräknar vi för varje händelse i vår historiska databas de korresponderade vädervariabler som gällde när händelsen inträffade. Det visas i Figur 3.2 var prickarna föreställer VViS-stationer och händelse inträffar på plats var nålen finns. Data från de 7 närmaste (röda) stationerna i den (röda) cirkeln som används för att beräkna variablerna på plats var händelsen inträffades, data från dem övriga (blåa) stationer används inte.

3.4 Datautforskning

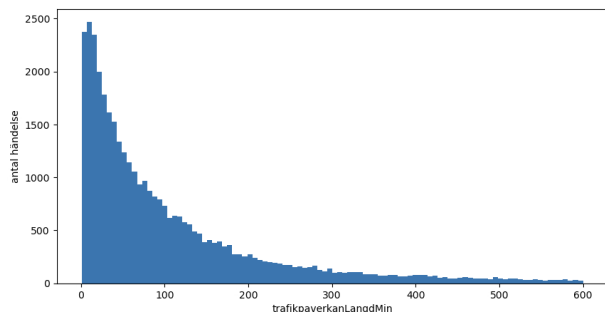
För att få en känsla för våra variabler utforskar vi dem först separat för att se hur dem är fördelat, om det finns felaktig data eller om data saknas. Figur 3.3 visar antal händelser per månad för hela dataset. De första 9 månader



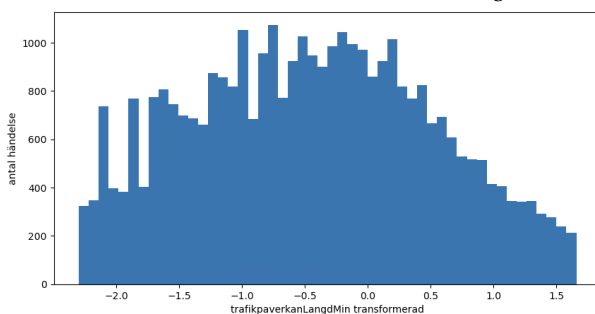
Figur 3.3: Antal händelser per månad.



(a) Samtliga händelser.



(b) Begränsad till och med 600 minuter



(c) Begränsad till och med 600 minuter och transformerad

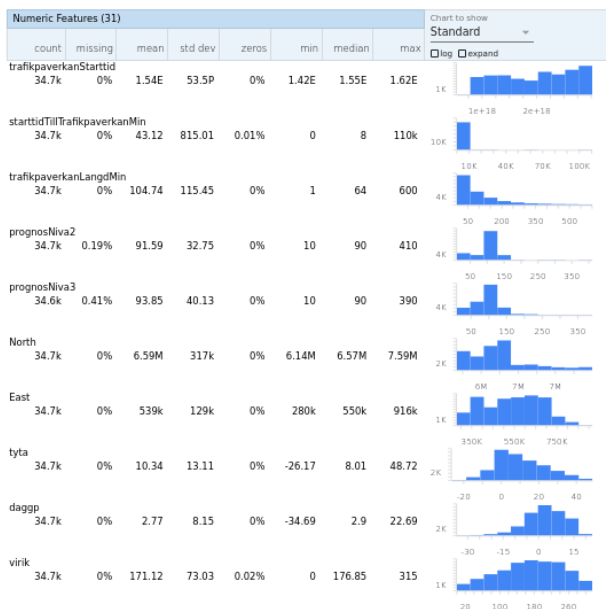
Figur 3.4: Distributioin av trafikpöverkanslängd.

av 2015 finns inte mycket användbar data, det är okänt var för data inte har samma kvalitet som senare åren då det är bättre. Tydligt är också att de flesta händelser är infrafel (örsakskodniva-I) och att driftfel som leder till påverkad trafik förekommer väldigt sällan.

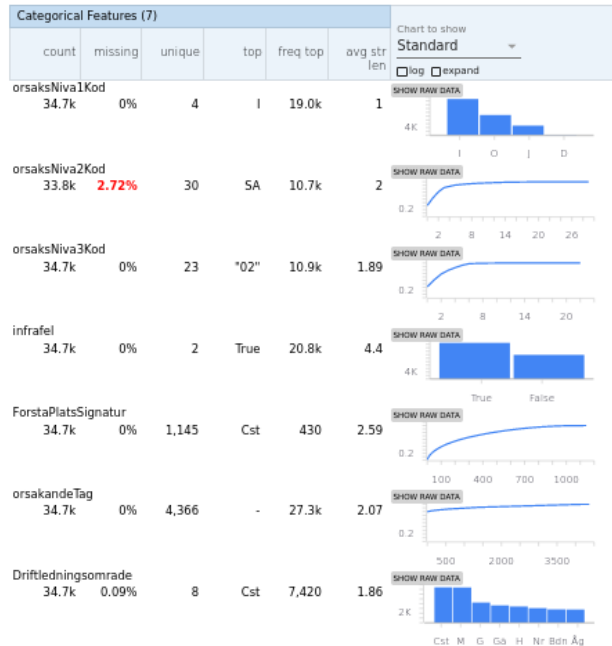
Figur 3.4 visar trafikpöverkans längd av hela dataset (a) och begränsad till och med 600 minuter (b). Vi ser målvariabeln är fördelat väldigt sned där de flesta observationer har ett lågt värde och få observationer som tar lång tid att lösa. Eftersom det är svårt för algoritmen att hitta mönster om det inte finns mycket data har vi beslutat att begränsa dataset till händelser som tar max 10 timmar innan trafiken körs som vanligt igen. Händelser som tar flera veckor eller längre är ofta också stora projekt som planeras in långt på förväg och därför inte behöver en prognos från algoritmen, så från en verksamhetsperspektiv passar filtreringen också bra. Vi också använder en transformering av målvariabeln, se figur 3.4(c) så att det blir mer normalfördelat och lättare för algoritmen att prediktera rätt. Vi använder **NumPy**'s *log1p* funktion som transformerar som följande:

$$x_{trans} = \log(1 + x) \quad (3.1)$$

Figur 3.5 visar en statistik sammanfattning av några av dem variabler som kommer från LUPP. Så ser vi till exempel att variabeln 'PrognosNiva2' saknas data för 0.19% av entries. Det betyder att vi behöver filtrera bort dessa rader eftersom vi vill jämföra våra prognoser med de historiska prognoserna.



(a) Numeriska variabler



(b) Kategoriska variabler

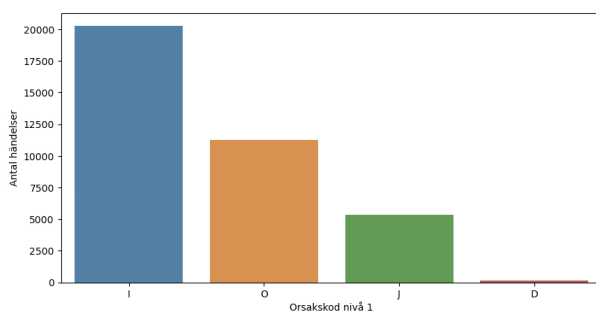
Figur 3.5: Översikt över variabler relaterade till händelser.

3.4.1 Kategoriska variabler

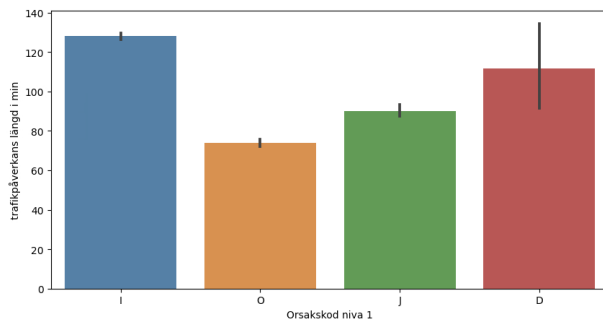
Här sammanfattar vi hur vi har utforskat de kategoriska variabler som finns och visar några resultat för att förtydliga. För att visualisera variabler använder vi så kallade stapeldiagram.

Figur 3.6(a) visar antalet händelser per orsakskodniva1 och Figur 3.6(b) visar snitt trafikpåkavens längd per orsakskodniva1. Som vi uppmärkte tidigare är dem flesta händelser av typ 'I', följd av 'O', 'J', och sist 'D'. Längden av trafikpåkavens visar att händelser av typ 'I' tar mest tid innan trafikpåkavens är löst, följd av 'D', 'J', och 'O'. Skillnaden är tillräckligt klar så att vi kan besluta att variabeln 'orsakskodniva1' kommer att vara en viktig variabel i vår modell. För att få djupare förståelse och mer förklaring av Orsakskoder vad de betyder, se [32].

Nu tittar vi lite mer granulärt på variabeln 'orsakskodniva2'. Figur 3.7(a) visar antalet sådana händelser per typ. Eftersom det är svårt att hitta mönster om det inte finns mycket data filtrerar vi bort händelser med orsakskodniva2 som förekommer mindre än 500 gånger. Det resulterar i de 9 orsakskoderna 'ISA', 'OMÄ', 'IBÖ', 'JDM', 'IEA', 'OTÄ', 'ONA', 'JVA' och 'OSY'. I figur 3.7(b) ser vi att de 9 typerna av händelser har stora skillnader i snitt trafikpåkavenslängd. Därmed bör orsakskodniva två vara viktig för modellen.

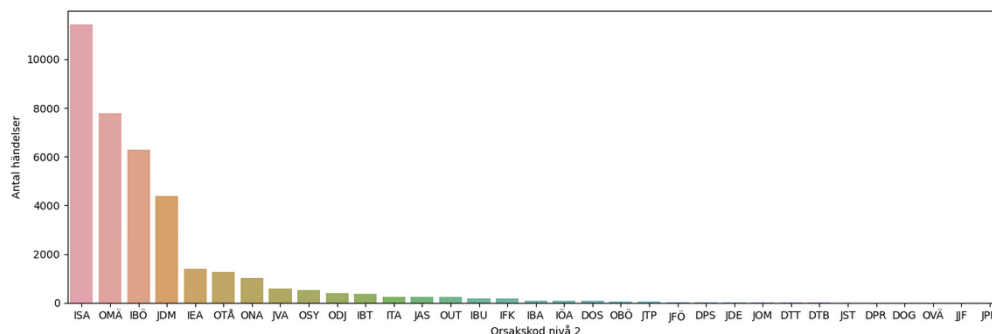


(a) Antalet händelser per orsakskodniva1

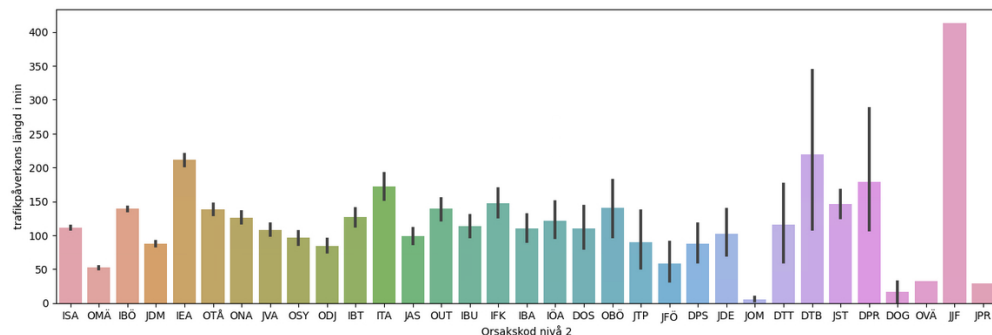


(b) Snitt trafikpåkavens längd

Figur 3.6: Statistik för 'orsakskodniva1'.



(a) Antalet händelser per orsakskodnivå 2



(b) Snitt trafikpåverkans längd per orsakskodnivå 2

Figur 3.7: Statistik för 'orsakskodnivå 2'.

Figur 3.8 visar antalet händelser uppdelat på orsaksnivåkod 2 över tid. Här kan vi få en översikt över mönster, så som att orsakskoder används olika mycket över tid eller att de verkar användas på ett annat sätt vid en viss tidpunkt. Vi vill säkerställa att datan inte har drastiska förändringar vilket kan påverka modellen. I det här fallet verkar det inte finnas några dramatiska ändringar. Om så hade varit fallet hade vi behövt undersöka andledningen till det och hur det bör hanteras utifrån ett verksamhetsperspektiv.

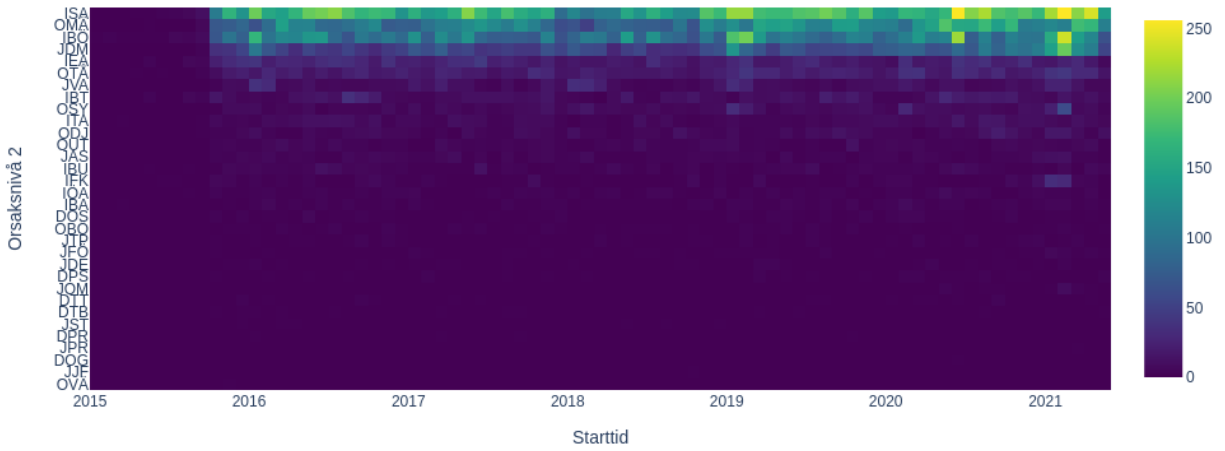
När vi utforskar 'Driftledningsområde' variabeln ser vi att flesta trafikpåverkande händelser förekommer i Stockholm och Malmö, men att snitt längden av händelser är störst i Boden, förmodligen eftersom avstånden, inställelse tid för entreprenörer och andra orsaker är mycket större där och det tar längre att ta sig till händelsens plats om nånting behöver åtgärdas, hämta material som exempel blir flera långa resor och där av tiden längre för händelsen.

Antalet händelser som inträffar i helgerna är tydligt mindre än på vardagar (Figur 3.10) men skillnaden i trafikpåverkans längd är inte statistiskt signifikant, så det finns ingen direkt koppling mellan vilken dag det är och hur länge det kommer att ta tills att trafiken är påverkad.

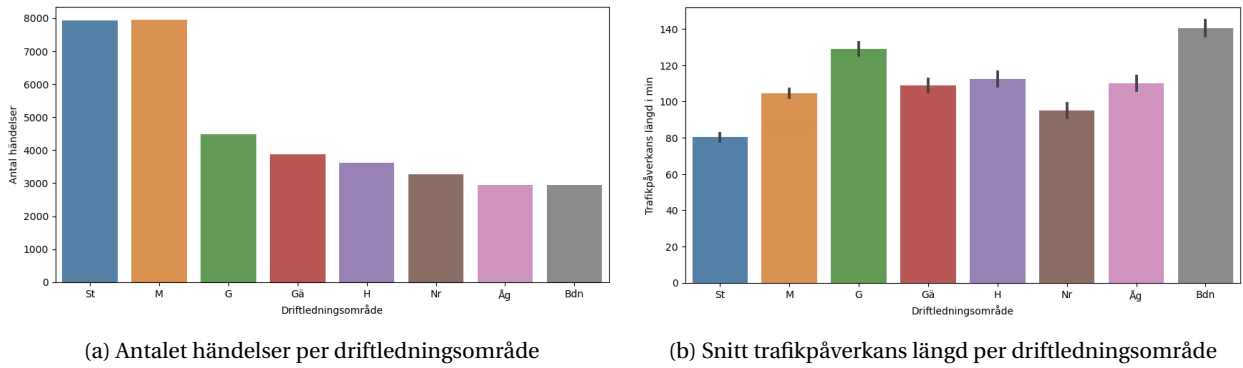
3.4.2 Kontinuerliga variabler

För att visualisera kontinuerliga variabler behöver vi andra typer av plottar. Vi använder histogram för att se fördelningen av mätningar, och scatterplot för att se eventuella samband mellan den kontinuerliga variabeln och målvariabeln. Till exempel tittar vi på 'tyta' som är mätningar på VViS stationer av yttemperaturen. I figur 3.11(a) ser vi att temperaturen ligger mellan -25 och 50 grader med dem flesta mätningar rund 0. I figur 3.11(b) visas att det finns ingen tydlig samband mellan yttemperaturen och händelsens längd. Om till exempel händelsen skulle ta längre om det är kallt skulle grafen visa ett mönster, men slumpmässigheten av mätningar är stor. Det kan naturligtvis alltid finnas mer komplexa samband, till exempel att 'orsakskodnivå1' tillsammans med 'tyta' påverkar händelsens längd. Men vår analys visar att det finns ingen vädervariabel som direkt påverkar längden av trafikstörningar vid en händelse.

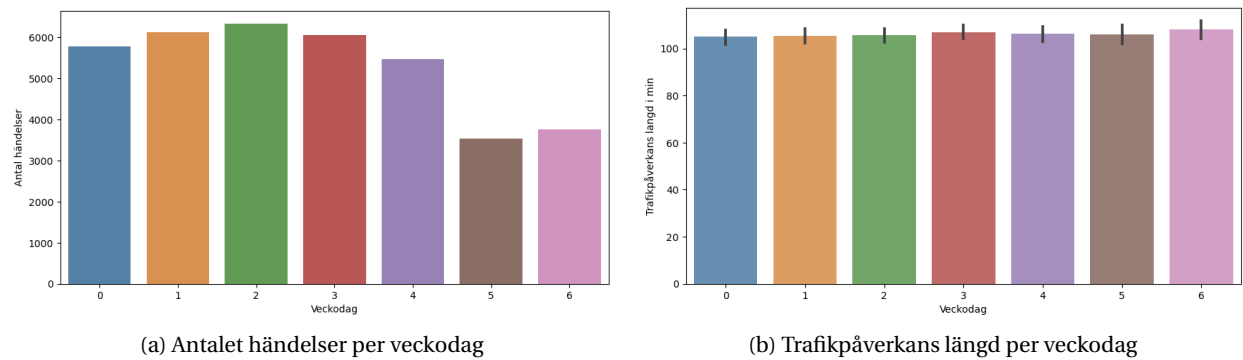
Förseningar (antal)



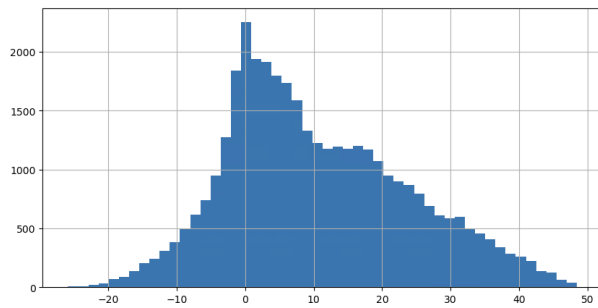
Figur 3.8: Antal trafikpåverkande händelser över tid för 'orsakskodniva2'.



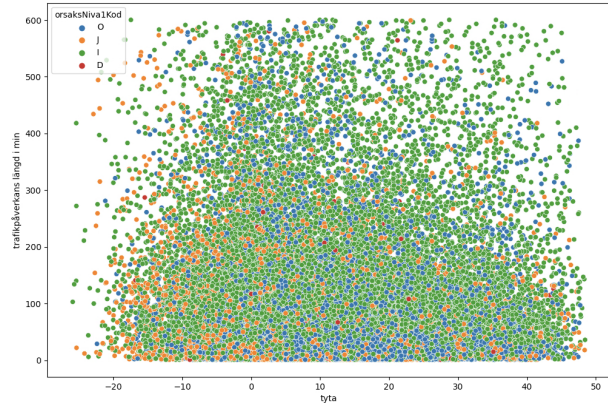
Figur 3.9: Statistik för 'driftledningsområde'.



Figur 3.10: Statistik för 'veckodag'.



(a) Histogram av 'tyta'



(b) Trafikpåverkans längd mot 'tyta'

Figur 3.11: Statistik för 'tyta'.

3.4.3 Interntext

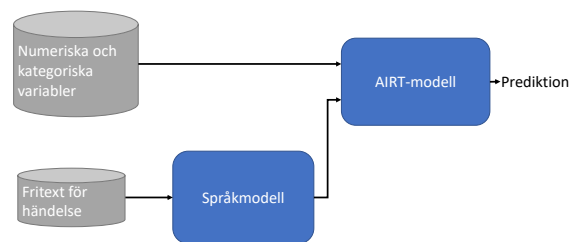
En trafikstörande händelse kan inkludera fritext som kan ge mer detaljer eller kontext till händelsens orsak. Ett exempel ges i figur 3.12 där ordet 'obehöriga' i händelsens text indikerar att trafikpåverkan vara kortare tid. Eftersom textfält kan innehålla information som inte återspeglas i övriga variabler bör de kunna ge modellen bättre förutsättningar att kunna förutspå trafikstörnings längden för en händelse.

En nackdel med att titta på enskilda ord är att ett ord kan betyda olika saker beroende på sammanhanget, det vill säga de andra orden i texten. För att ta hänsyn till hela texten kan vi använda oss av språkmodeller som är tränade på stora textmängder för att lära sig både meningsuppbyggnad och ords betydelse i förhållande till varandra.

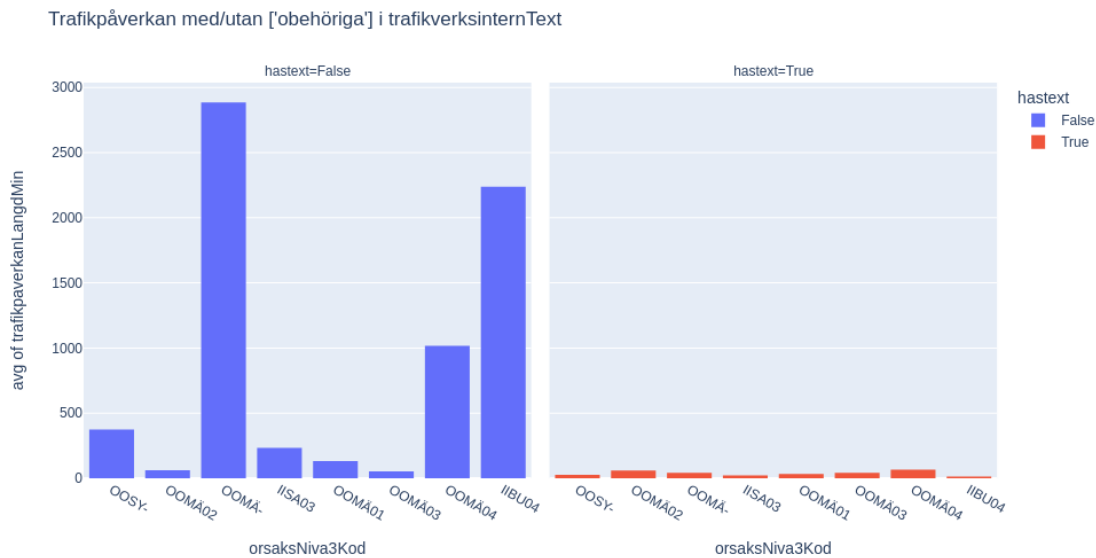
Modeller för naturligt språk tolkar ord och meningsbyggnad. De kan till exempel tränas för att översätta text eller förutspå nästkommande ord. För att göra dessa uppgifter lär sig modellen vilka meningar som betyder liknande saker och producerar numeriska värden som beskriver var de befinner sig i 'meningsrymden'. Detta faktum har vi utnyttjat i det här projektet genom att använda de numeriska värdena som språkmodellen producerar för att ge extra information till vår modell som inte hanterar text-data, se figur 3.13.

En populär språkmodell är BERT som även finns för svenska språket. Vi har därför använt oss av en svensk BERT-modell för att omvandla texten till en numerisk representation som kan användas för prediktion av trafikstörningslängd. För att förstå om språkmodellen förstår interntexten för händelserna har vi, för första rapporten i varje händelse, kört texten genom språkmodellen. För varje händelse har vi därför en högdimensionell vektor som beskriver texten.

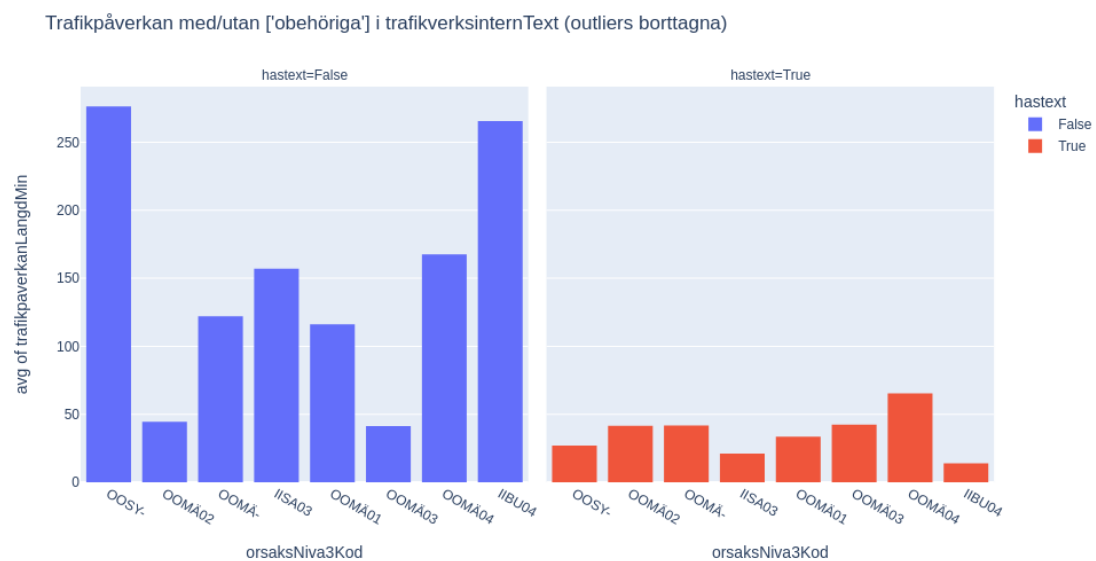
Eftersom det är svårt att granska en så pass högdimensionell vektor (500-1000 dimensioner) har vi använt oss av en projektmetsod (UMAP) som försöker behålla avstånden mellan vektorerna i den högdimensionella rymden i två dimensioner. Resultatet av projektionen kan ses i figur 3.14 där varje prick representerar en händelses interntext och färgen representerar trafikpåverkanslängden. Vi kan se att modellen skiljer på till exempel texter utan innehåll, med många punkter, och även vad de handlar om till viss utsträckning (brand osv.). Vi kan även se att det inte finns självklara områden där trafikpåverkan skiljer sig mycket från övriga områden eftersom färgerna är utspridda till synes slumpmässigt. Detta indikerar att språkmodellens tolkning av interntexten kan vara svår för en modell



Figur 3.13: Översikt över hur text, numerisk samt kategorisk data processades av språk- och neuralt nätverk för att prediktera trafikpåverkanslängd.

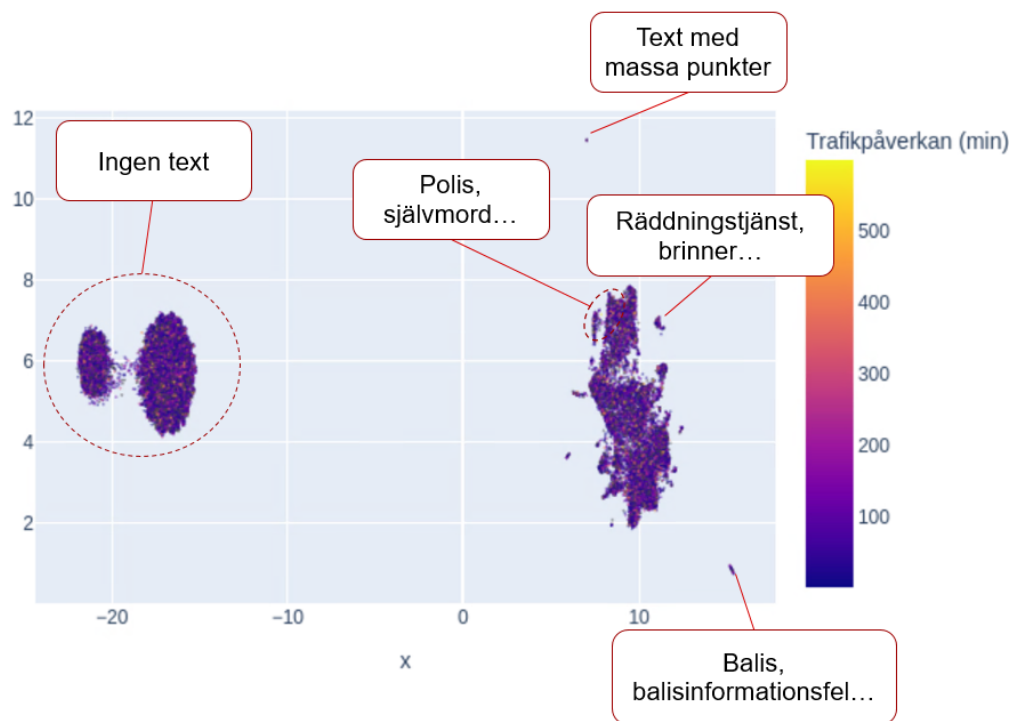


(a) All data.



(b) Utan extremvärden.

Figur 3.12: Relation mellan händelser som har/inte har interntext för händelser med orsakskod "obehöriga". I (b) har extremvärden tagits bort för att tydliggöra skillnaden mellan de händelser som har texten och de som inte har. För just den här orsakskoden verkar det finnas en större osäkerhet kring trafikpåverkan när inget skrivits kring händelsen.



Figur 3.14: Projektion av språkmodellens numeriska tolkning av interntext. Varje prick motsvarar interntext för en händelse och dess färg motsvarar trafikpåverkanslängden. Ju mer liknande färger är grupperade desto enklare är det för modellen att utnyttja informationen. Även om texter med liknande betydelse finns nära varandra så är tyvärr händelser med hög trafikpåverkan utspridda, vilket indikerar att det blir svårt för modellen att utnyttja informationen.

att använda för att förutspå trafikpåverkan.

Då vi kombinerade interntexten med övrig data enligt figur 3.13 så fick vi liknande resultat som utan interntexten. Vi kan därmed dra slutsatsen att antingen har orsakskoderna fångat upp betydelsen av interntexten eller så behövs en språkmodell som bättre tar tillvara på Trafikverkets domänspecifika språk (exempelvis balis, växelfel osv). Då användandet av interntexten inte bidrog till att öka prestandan har vi utelämnat den i det fortsatta arbetet.

3.5 Modellering

Olika typer av modell testades:

- Linear Regression
- Ridge Regression [33]
- Lasso Regression [34]
- Random Forest [35]
- Gradient Boosting [36]
- AdaBoost [37]
- Neuralt nätverk [38]

Vi genomförde en enkel hyperparameter-optimering, men resultaten mellan olika modell och parameter inställningar var ganska lika så att vi beslutade att inte spendera mer tid på en djupdykning för att undersöka bäst möjlig parameterset.

3.6 Resultat

För att undersöka vilka parameter kan hjälpa att förutspå målvariabeln ('trafikpaverkanLangdMin') undersöktes vi först datan (se sektion 3.4). I själva modellbyggande började vi med en modell med detsamma inputvariabler som Trafikverket använder just nu för att beräkna genomsnittliga tiden av en händelse ('orsaksNiva1Kod', 'orsaksNiva2Kod', 'orsaksNiva3Kod', och 'Driftledningsomrade'). Vi utökade data settet med en ny variabel för att se om det förbättrade resultatet. Om det fanns en signifikant förbättring behöll vi nya variabeln, om inte tog vi inte med variabeln i modellen.

Vi experimenterade även med att använda en svensk variant av språkmodellen BERT för att få in intertext-variabeln i modellen (se sektion 3.4.3). Intertextfältet kan innehålla information som inte framgår av övriga variabler. Våra experiment visade dock att intertexten inte förbättrade modellen. Detta kan bero på att språkmodellen BERT inte har förståelse för järnvägs-koncept så som *växelfel*.

Till slutet var det följande variabler som visades påverkan längden av en händelse:

- *orsaksNiva1Kod*
- *orsaksNiva2Kod*
- *orsaksNiva3Kod*
- *Driftledningsomrade*
- *infracfel*
- *inställelsetid*
- *driftNiva_max*.

Första fyra variabler är självförklarande, 'infracfel' är en boolean som beskriver om det finns en infracfelsrapport vid första anmälan i systemet eller ej, 'inställelsetid' beskriver hur länge en operatör har för att vara på plats för att sedan undersöka en händelse enligt baskontraktet de har med Trafikverket, och *driftNiva_max* beskriver max-nivån av driftnivån under en händelsens livstid (från 0 som betyder ingen trafikpåverkan alls, till 4 komplett stop.).

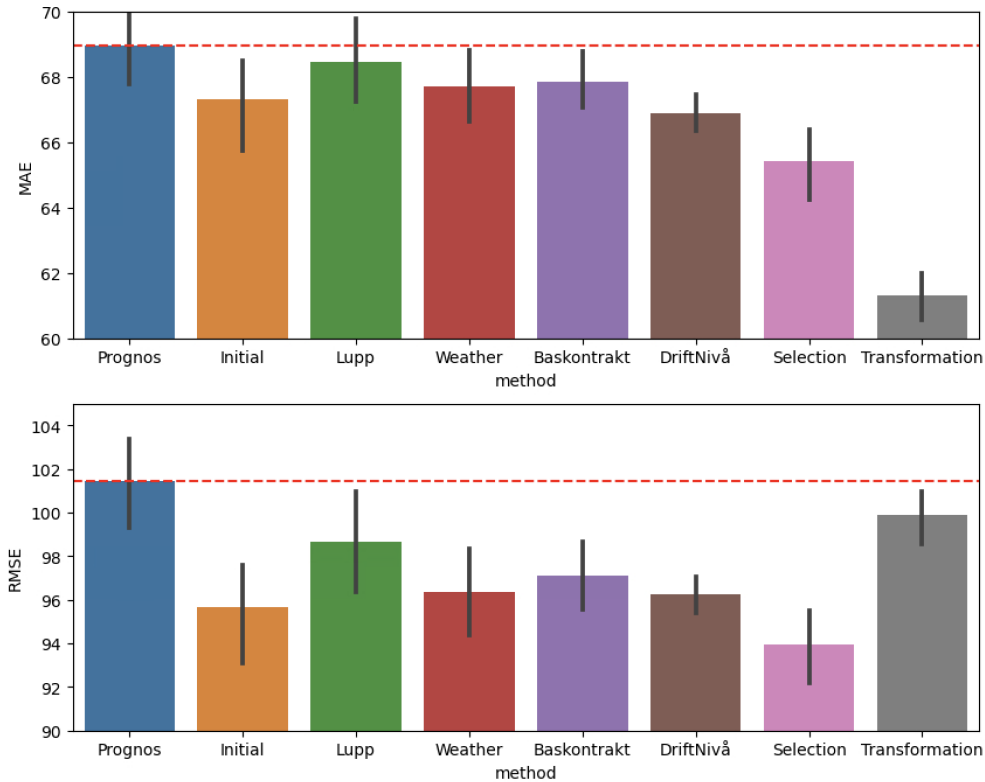
För att utvärdera modellerna behöver vi definiera ett fel eller loss funktion. Vi valde att algoritmen minimerar Mean Absolut Error (MAE), men vi också kollade Root Mean Square Error (RMSE), definierad som:

$$MAE = \frac{\sum_{i=1}^N ||y_i - \hat{y}_i||}{N} \quad (3.2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N ||y_i - \hat{y}_i||^2}{N}} \quad (3.3)$$

med y_i faktiska försening av händelse i , och \hat{y}_i modellens prediktion. Båda fel-måtten är vanliga i maskinlärning som utvärdering, men observera att en egenskap av RMSE är att den är mer känslig för extremvärden än MAE.

Figur 3.15 visar resultaten av modellerna. Alla modeller presterar bättre än nuvarande prognoser när vi jämför MAE, men att lägga till vädervariabler ('Weather') och/eller baskontraktdata förbättrar modellen inte signifikant. Att lägga till information om maximum driftnivån ('DriftNivå'), och använda bara de typer av händelser med mer än 500 händelser i historiska datasetet ('Selection') förbättrar resultatet. Observera att det finns en stor förbättring vid transformering av målvariabeln i MAE, men att RMSE går upp. Anledningen till detta ligger i hur modellen hanterar extremvärden. Efter transformation är all data mer koncentrerad och finns mindre extremvärden, då minimerar algoritmen MAE. För att beräkna felet borde vi transformera tillbaka till originalet, och då är RMSE mer känslig för extremvärden. Man kan också dra slutsatsen att algoritmen borde minimera RMSE men det beslutet är beroende på vad är viktigt för verksamheten.



Figur 3.15: MAE och RMSE i minuter för olika steg i modelleringen.

Figur 3.16 visar på hur mycket olika variabler påverkar modellens prediktion. SHAP fungerar på så sätt att man tränar många modeller och ser hur prediktionen skiljer sig från medelvärdet när man tar bort en variabel. Vi kan se att orsakskoder påverkar prediktionen mest. En lång starttid till trafikpåverkan (starttidTillTrafikpaverkanMin, röd färg), alltså tiden från att rapporten skrivs till att den är trafikpåverkande, gör även att modellen tror att trafikpåverkan blir längre. En hypotes är att till exempel planerat underhåll kan pågå under längre tid och att dessa typer av händelser skapas lång tid innan trafikpåverkan tar effekt.

Väder, som till exempel ned_typ (nederbördstyp) och tyta (yttemperatur), påverkar trafikpåverkanslängden något men inte mycket.

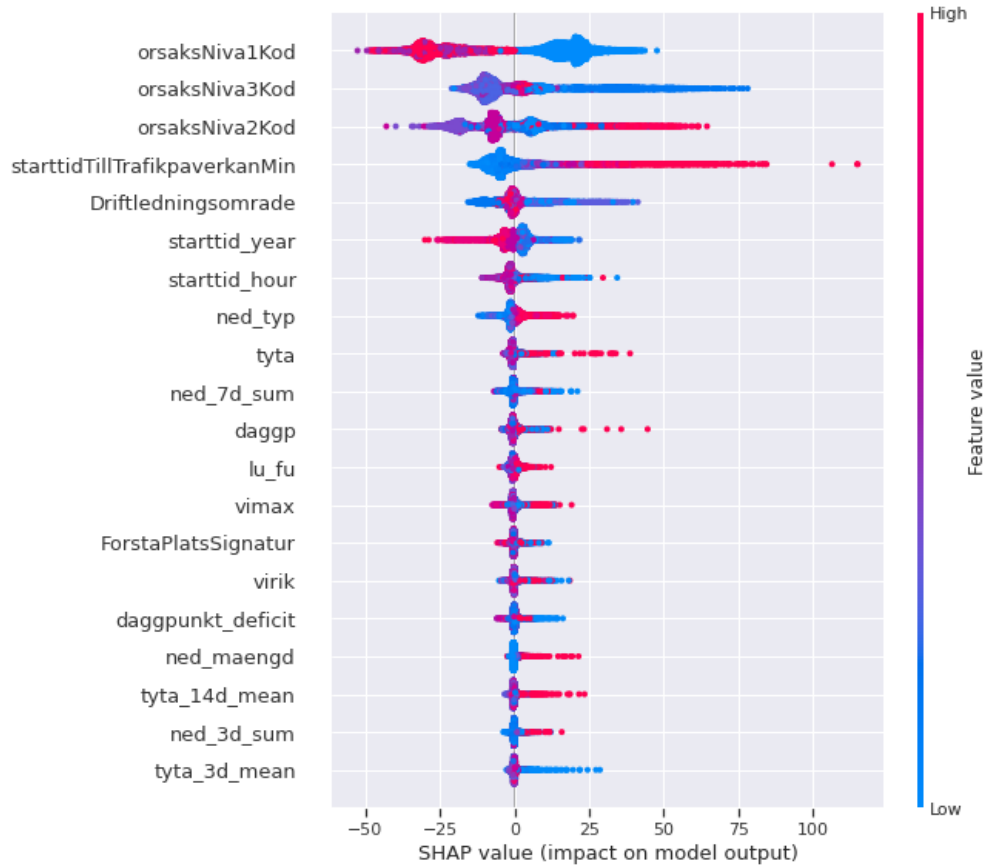
3.7 Diskussion och slutsatser

Variabeln trafikpåverkan är problematisk av flera orsaker. För det första beror den på när operatören har tid att rapportera början och slut på trafikpåverkan. För det andra beror den på hur trafiken ser ut just vid tillfället för händelsen. Samma typ av händelse som tar precis lika lång tid att åtgärda kan alltså ha olika trafikpåverkan ifall det går tåg på sträckan eller inte.

Åtgärder för att förbättra situationen för trafikpåverkan-variabeln skulle kunna vara att 1) inkludera automatiserad analys av när tågtrafiken börjar och slutar påverkas av händelsen samt 2) inkludera hur trafiken ser ut för associerade sträckor.

Med detta i åtanke kan vi ändå se att en mer avancerad modellbaserad prediktion är bättre än en statistikmodell baserad på orsakskod och driftområde. Störst skillnad i resultat för MAE fås då vi gör distributionen för variabeln 'trafikpåverkan' är mer lik en normalfördelning. I det fallet är den mer avancerade modellen i snitt ca 7 minuter (11%) bättre än prognos fig 3:15 som representerar den manuella förmågan till prognos.

Ett annat problem är att det finns inte mycket data för flera orsakskoder. Det är svårt att upptäcka mönster om det finns brist av data. Vi försökte motverka det med att producera syntetisk data, men det gav inte förväntad resultat. Att producera syntetisk data från ett existerande dataset är ett ungt forskningsområde men det visar mycket



Figur 3.16: SHAP-analys av variablers påverkan på en modells prediktion. Variablerna högst upp har störst påverkan på prediktionen och färgen anger om värdet på en variabel är högt eller lågt.

potentiell.

Kapitel 4

Ankomsttid för tåg

4.1 Introduktion

Prognoser kring när ett tåg anländer till station är en viktig del av trafiknätets dagliga verksamhet, dels för resenärer men också trafikledare. Prognoser gällande kommande förseningar omfattar ett komplext system med flera påverkande faktorer så som tågets tidigare förseningar, sträckan och stationers utformning, tid och plats tillsammans med andra tåg på och på kringliggande sträckors påverkan på tåget i fråga.

Inget av systemen som används för prognostisering av ankomst och avgång idag använder sig av historisk data och maskininlärning för att generera prognoser. Genom att använda historisk data kan vi ta fram automatiserade förseningsprognoser som baserar sig på mönsterigenkänning kring hur läget ser ut och hur det historiskt sett ter sig i samma situation.

4.2 Datautforskning

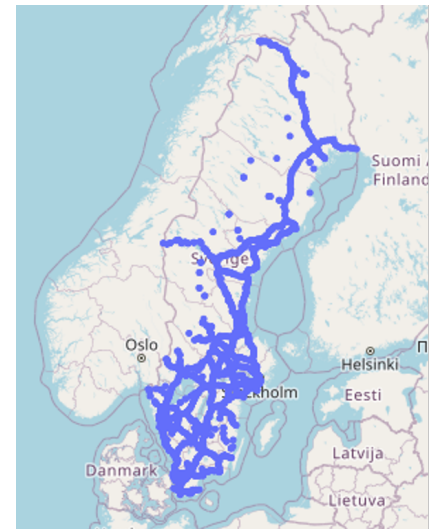
Den data som användes för ändamålet omfattar tidsseriedata gällande ankomster och avgångar för tåg i Sveriges tågnät under perioden **2021-11-01** och **2022-11-15**. Över den utvalda tidsperioden som återfinns datapunkter för samtliga platser i Sverige som visas i Figur 4.1. Datapunkter för ett unikt tåg delar uppdragsnummer och utgångsdatum. Längs den sträcka som tåget färdas stannar tåget vid eller åker förbi olika platssignaturer (eg. stationer). Ordningen för när dessa platssignaturer passeras definieras av platsens sekvensnummer för just detta tåg. Genom denna information går det att härröra sträckans startstation och ändstation, vilket tillåter att filtrera på sträckor mellan två stationer.

Vidare har endast försenade tåg beaktats i detta projekt och endast resenärståg har utvärderats och tagits med i test-settet. Eftersom godståg och andra tågslag kan påverka resenärståg har dessa tagits med i träningsprocessen av modellerna.

Majoriteten av platssignaturer längs sträckorna är endast passager och innebär inte ett stopp. Därför är oftast ankomsttiden och avgångstiden densamma för dessa platssignaturer. Om ett tåg stannar vid en platssignatur eller inte är också definierat i den data som används.

Varje tåg i Sveriges tågnät har specifika attribut som påverkar dess förutsättningar att anlända och avgå i tid samt dess förutsättningar att ta igen en försening. Dessa attribut tar form i exempelvis tågets tidigare förseningar, tågvikt, antal vagnar, drivmedel eller trafikutövare.

Det absolut viktigaste attributet för prognostisering av tågets ankomst och avgång till station är dess tidigare förseningar över sträckan. Ett försenat tåg kommer att bli mer försenat. Likväl är det större sannolikhet för ett tåg som är försenat vid en station A_1 att också bli försenat vid efterföljande station A_2 än att ankomma i tid. Därför är en



Figur 4.1: Karta över mätpunkter där tågen passerar.

stor del av framtida prognoser att se hur den senaste förseningstrenden ser ut.

Utöver tågets egna attribut så spelar sträckan som tåget befinner sig på en stor roll. Sträckan i fråga omfattar även de tåg som befinner sig på sträckan eller på angränsande sträckor.

Tågsystemet är ett komplext system och aktuella trafikmönster påverkar enskilda tågs förseningar. Därför är det viktigt att ta hänsyn till andra tågs situationer vid prognostisering av det enskilda tåget. Hur komplext man utvecklar nätverkets attribut är direkt begränsat till den typ av maskininlärningsmodellering man använder sig av.

4.3 Modellering

Prognostiseringen genomfördes och testades med diverse maskininlärningsmetoder. Modellerna som testades var primärt trädmodeller:

- XGBoost Regression (Extreme Gradient Boosted Regression)
- LGBM Regression (Light Gradient Boosting Machine Regression)

Output för vardera modell var ankomstförsening (minuter) och avgångsförsening (minuter). Alla modeller tillåter inte två output genom regression varpå vissa av modellerna delades upp i två modeller - en tränad för ankomstprognos, och en tränad för avgångsprognos. Modeller likt XGBoost tillåter regression med två outputs, varpå denna modell inte behövde delas upp.

Baselines som använts för jämförelse i prestanda är tidtabellen och de manuella prognoser som finns tillgängliga i beräkningsrapporter från LUPP. Tidtabellen kan anses vara en modell som alltid predikterar 0 förseningsminuter, eg. att tåget alltid är i tid och följer tågtidtabellen.

De manuella prognoser som finns tillgängliga i LUPP är skapade av trafikledare på trafikcentralen runt om i Sverige. Trafikledare sitter inne på stort erfarenhetskapital och stor igenkänningsförmåga inom järnväg, varpå de prognoser som sätts, sätts med hjälp av mänsklig förmåga att förutse framtida utfall. Enligt de produktionsledare som intervjuats i samband med detta projekt så kommer den stora erfarenheten genom försök som lyckats och misslyckats vad som är bäst.

4.3.1 Features och Hyperparameterisering

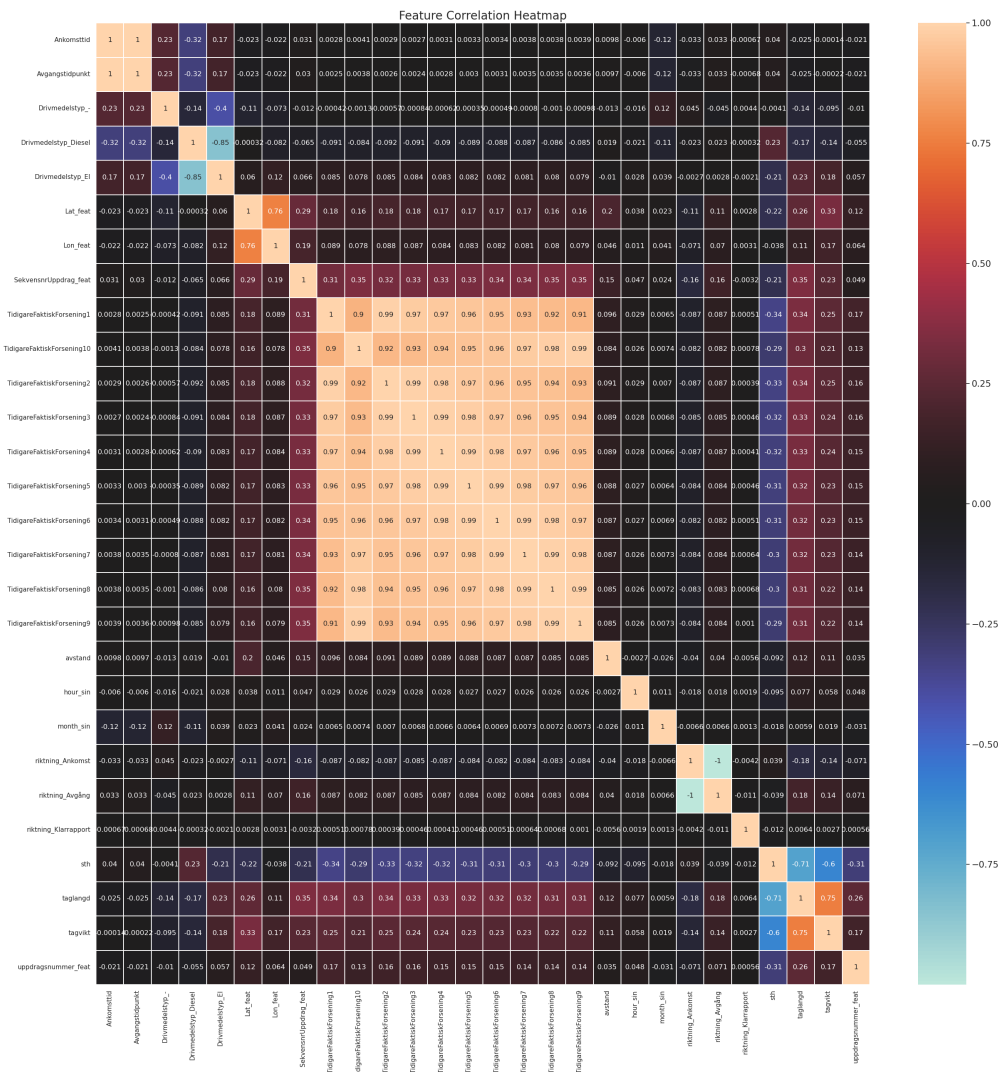
Modellerna beskriva i avsnitt 4.3 använde sig av flertalet features (input) för att sätta prognoser gällande tidsavvikelser i minuter för ankomst och avgång för vardera station över en sträcka. Samtliga features skalades, antingen numerärt, periodiskt eller kategoriskt. Detta för att underlätta för modellerna i mönsterigenkänningen och prognostiseringen.

I avsnitt 4.2 beskrivs bland annat hur features kopplade till det enskilda tåget använts. Dessa features är tågets tidigare avgångsförseningar, dess drivmedelstyp, vikt och tåglängd.

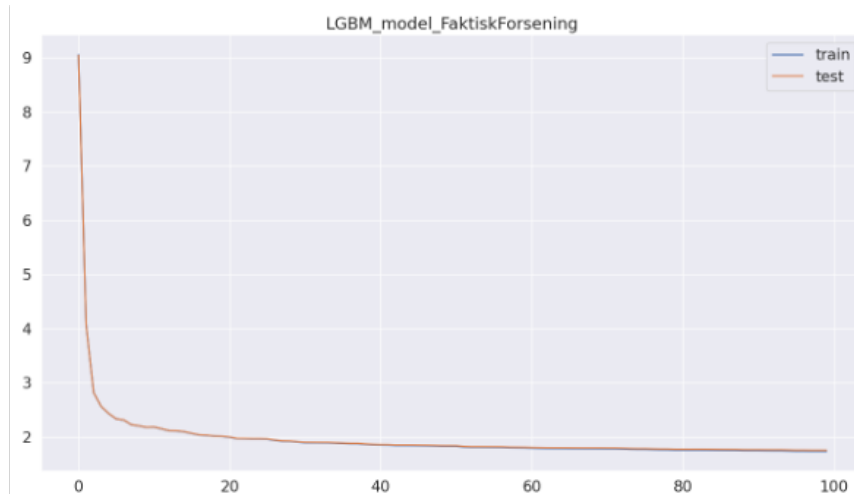
Utöver dessa features används också features kopplade till tågtidtabellen och sträckan. Dessa är tidtabell för ankomst och avgång, avstånd till nästkommande platssignatur, om tåget ska stanna eller inte, eventuella hastighetsnedsättningar på sträckan, timme på dygnet och månad på året för tågets nuvarande plats.

De features som använts i detta projekt förhåller sig till varandra enligt korrelationsvärmekartan i Figur 4.2. I denna korrelationsvärmekarta beskriver rader och kolumner de features som använts och varje cell beskriver korrelationen mellan dessa. Diagonalen beskriver varje features korrelation till sig själv, varpå denna visar 1 (100%) i korrelation. I denna värmekarta kan vi se att alla features förutom de kategoriska features som används (eg. riktning och drivmedelstyp) samt tidigare förseningar har en låg korrelation, vilket är att önska då vi inte vill ha onödigt information i vårt dataset. Features som beskriver tidigare förseningar är djupt korrelerade vilket syns i korrelationsvärmekartan. Anledningen varför dessa fortfarande är kvar är för att påvisa den senaste trenden för respektive tåg, som återfinns i dessa features. Datasettet som sträckte sig över 1 års tidsseriedata delade upp i ett träningsset och ett testset enligt 90-10%, kronologiskt uppdelat. Detta innebar att den tidsmässigt senaste datan användes som testdata.

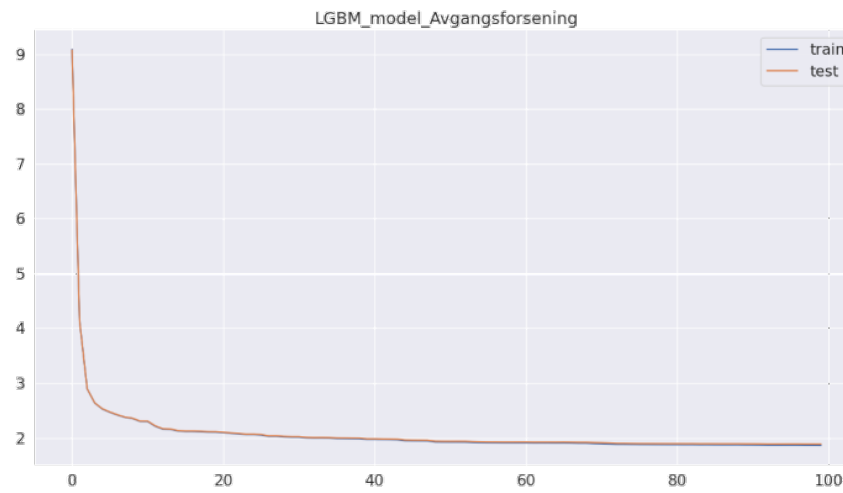
Modellerna tränades med hyperparameterisering. Träningen använde sig av loss function mean absolute error över 100 epoker och, enligt Figur 4.3 visade inga tecken på överfitning.



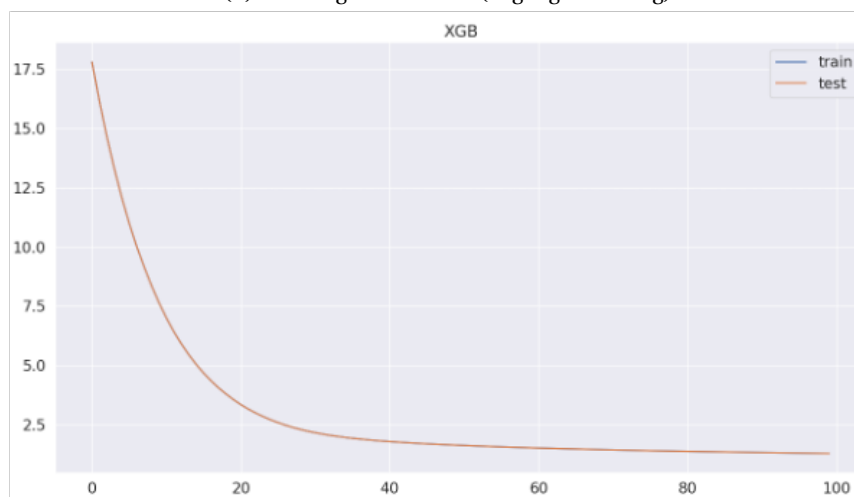
Figur 4.2: Feature korrelationsvärme-karta. Rader och kolumner beskriver features och varje cell beskriver korrelationen mellan dessa. Diagonalen beskriver en features korrelation till sig själv, varpå denna visar 1 (100%) korrelation.



(a) Learning curve LGBM (Ankomstförsening)



(b) Learning curve LGBM (Avgångsförsening)



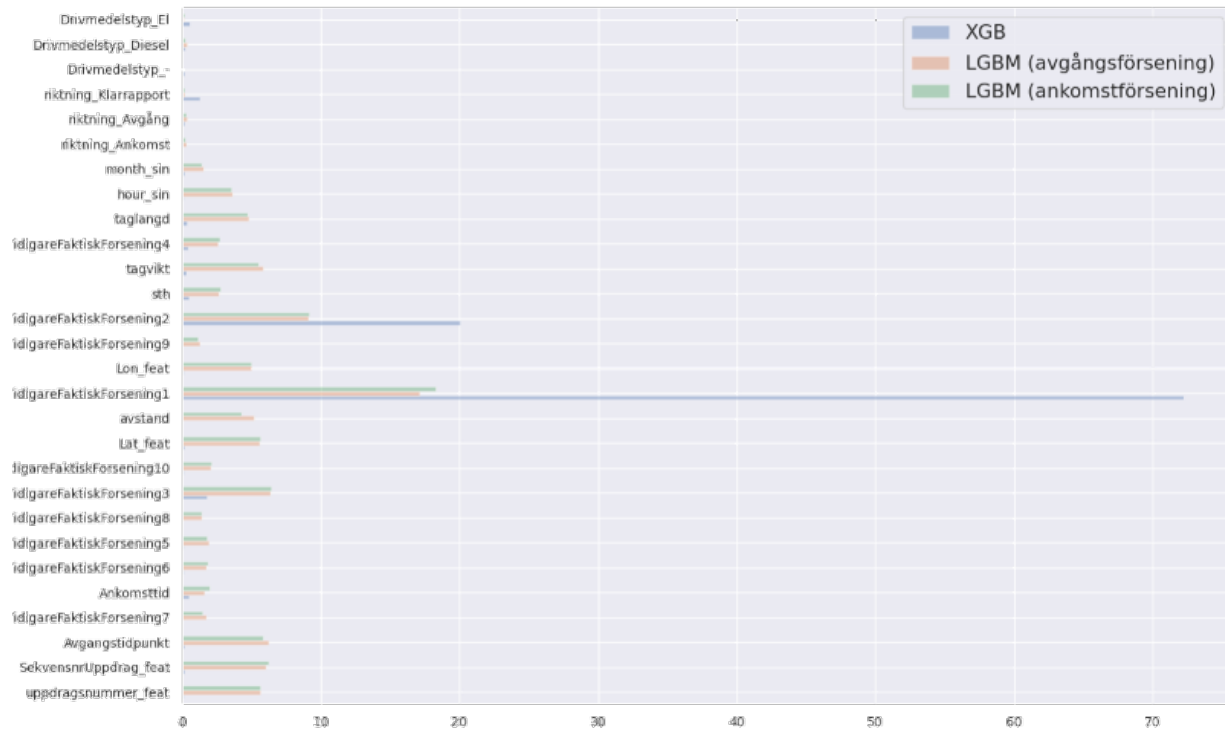
(c) Learning curve XGB (Ankomstförsening & Avgångsförsening)

Figur 4.3: Learning curve för de tre modeller som tränades. Som nämnt i avsnitt 4.3, tränades två LGBM-modeller - en för varje målvariabel (ankomst -och avgångsförsening).

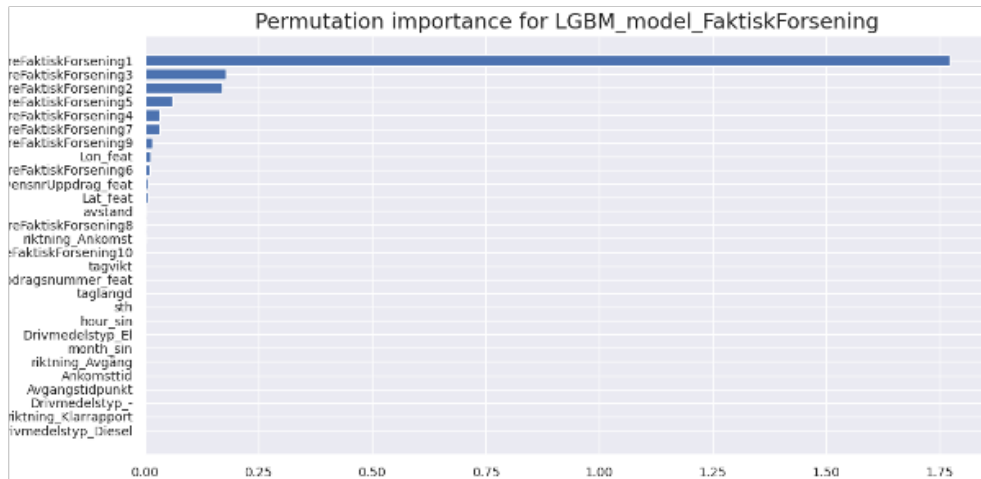
4.3.2 Feature Importance och modellförklaring

I trädmodeller är det av vikt att förstå hur modellerna, och dess output, i fråga fungerar och påverkas av de features som används som input. Feature importance är ett bra narrativ för detta ändamål då det ger en bra överblick kring magnituden av påverkan vardera feature har på modellen.

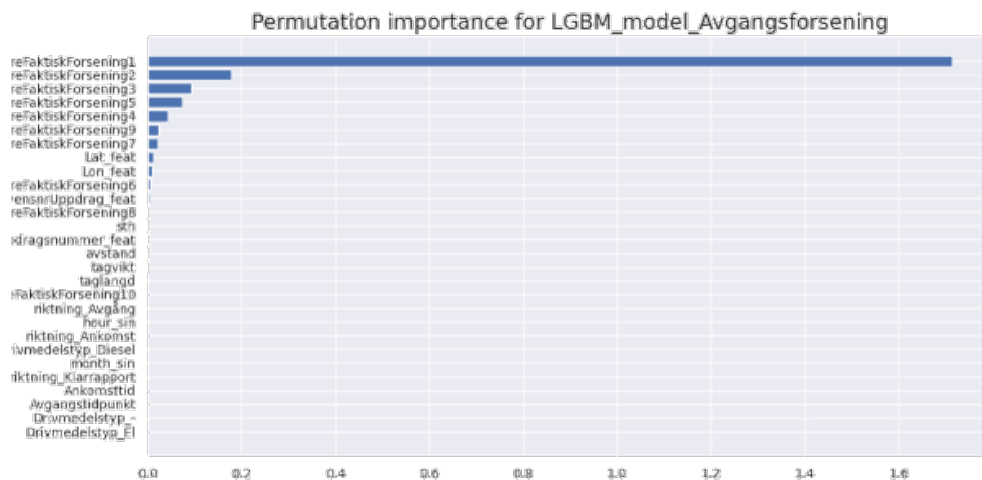
Feature importance kan presenteras och representeras på flera olika sätt. Av de metoder som går att använda har vi valt att presentera modellernas inbyggda feature importance (se figur 4.4) samt permutation importance (se figur). Vad som kan urskiljas ur modellernas feature -och permutation importance är att den i särklass viktigaste parametern gällande avgång och ankomst är vad den tidigare förseningen var, vid stationen innan. Vi kan också se att LGBM-modellerna har en jämnare fördelning av feature importance än vad XGB-modellen påvisar. LGBM-modellerna visar upp en tydligare användning och påverkan av alla de features som faktiskt används i modellen.



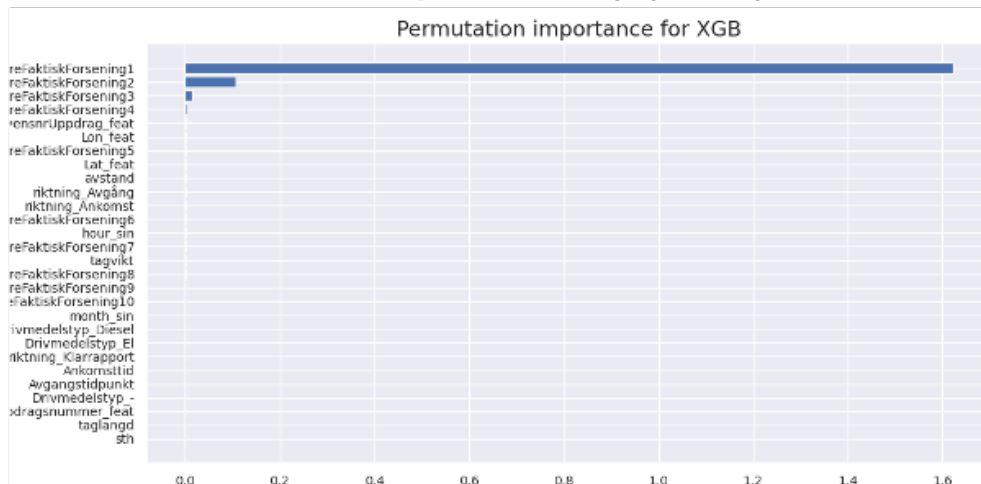
Figur 4.4: Feature Importance för de tre modellerna



(a) Permutation importance LGBM (Ankomstförsening)



(b) Permutation importance LGBM (Avgångsförsening)



(c) Permutation importance XGB (Ankomstförsening & Avgångsförsening)

Figur 4.5: Permutation importance för de tre modeller som tränades. Som nämnt i avsnitt 4.3, tränades två LGBM-modeller - en för varje målvariabel (ankomst -och avgångsförsening).

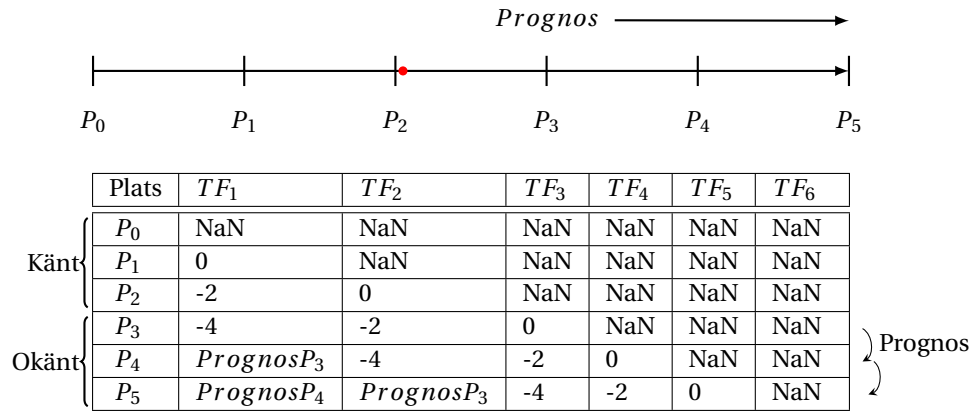
4.3.3 Framtida prognoser

Prognostisering gällande ankomst och avgång är inte bara intressant för nästkommande station, utan alla framtida stationer och stopp på sträckan. Därmed blir framtida prognoser en slags tillståndsformulering baserat på tidigare prognoser. Detta orsakar att desto längre fram på sträckan en prognos efterfrågas, desto större felmarginal kommer prognosen att ha. Eftersom tidigare förseningar omfattar en stor innebörd för prognostisering av förseningar så blir utmaningen med prognostiseringen desto svårare.

Tillståndsalgoritmen utvecklad för framtida ankomster och avgångar fungerar genom att sekvensiellt använda prognoser som input för tidigare förseningar på sträckan enligt ekvation 4.1.

$$PrognosP_m = f(P_{m-1}, x_{m-1}) = f(P_{m-1}, g(x_{m-1}, x_{m-1})) \quad (4.1)$$

där f syftar till funktion gällande prognostisering och g syftar till funktion gällande tillståndsförändring. Praktiskt sett, påverkar detta vår input till vardera plats enligt Tabell 2



Tabell 4.1: Exempel och förklaring hur tillstånd förändras för framtida prognostisering baserat på tidigare prognoser. Övre linjen beskriver en sträcka mellan plats P_0 och P_5 varpå tåget i fråga befinner sig vid röd punkt vid plats P_2 och prognoser för ankomst och avgång efterfrågas för plats P_3 , P_4 och P_5 . Tabellen visar n senaste förseningar (TF_n) för varje plats P_m . Prognoser för plats P_3 , P_4 och P_5 använder prognoser för tidigare platser.

4.4 Resultat

Hur bra en prognos är i tågnätet, kan beskrivas på flera olika sätt. Dels kan man kartlägga hur bra prognosen är för nästkommande stations eller för flertalet stationer framåt i tiden. Vidare, framtiden i tågsystemet kan beskrivas på flera olika sätt - tidsmässig framtid (ex. prognoser sätts 1h innan tabellenlig ankomst), stationsmässig framtid (ex. prognoser sätts 5 stationer innan tabellenlig ankomst) eller distansmässig framtid (ex. prognoser sätts 100km innan tabellenlig ankomst). För samtliga prestandamått har projektets baselines prestandamått tagits fram. Tidsperioden för evalueringsdatasetet täcker **2022-10-01 - 2022-11-15**.

I Figur 4.6a ses vilka platser i evalueringssettet vilka platser i Sverige som manuella prognoser har gjorts och i Figur 4.6b ses vilka datapunkter som återfinns i evalueringssettet och de platser som prognoser utförts vid.



(a) Manuella prognoser



(b) LGBM, XGB, Tidtabell

Figur 4.6: Platser i evalueringssettet där prognoser genomförs. 4.6a visar de platser där manuella prognoser genomförs och 4.6b visar vilka platser där AI-modeller (XGB och LGBM) samt tidtabellen har prognoser. Platser skiljer sig från 4.1 då tidsperioden för 4.6b täcker en kortare period.

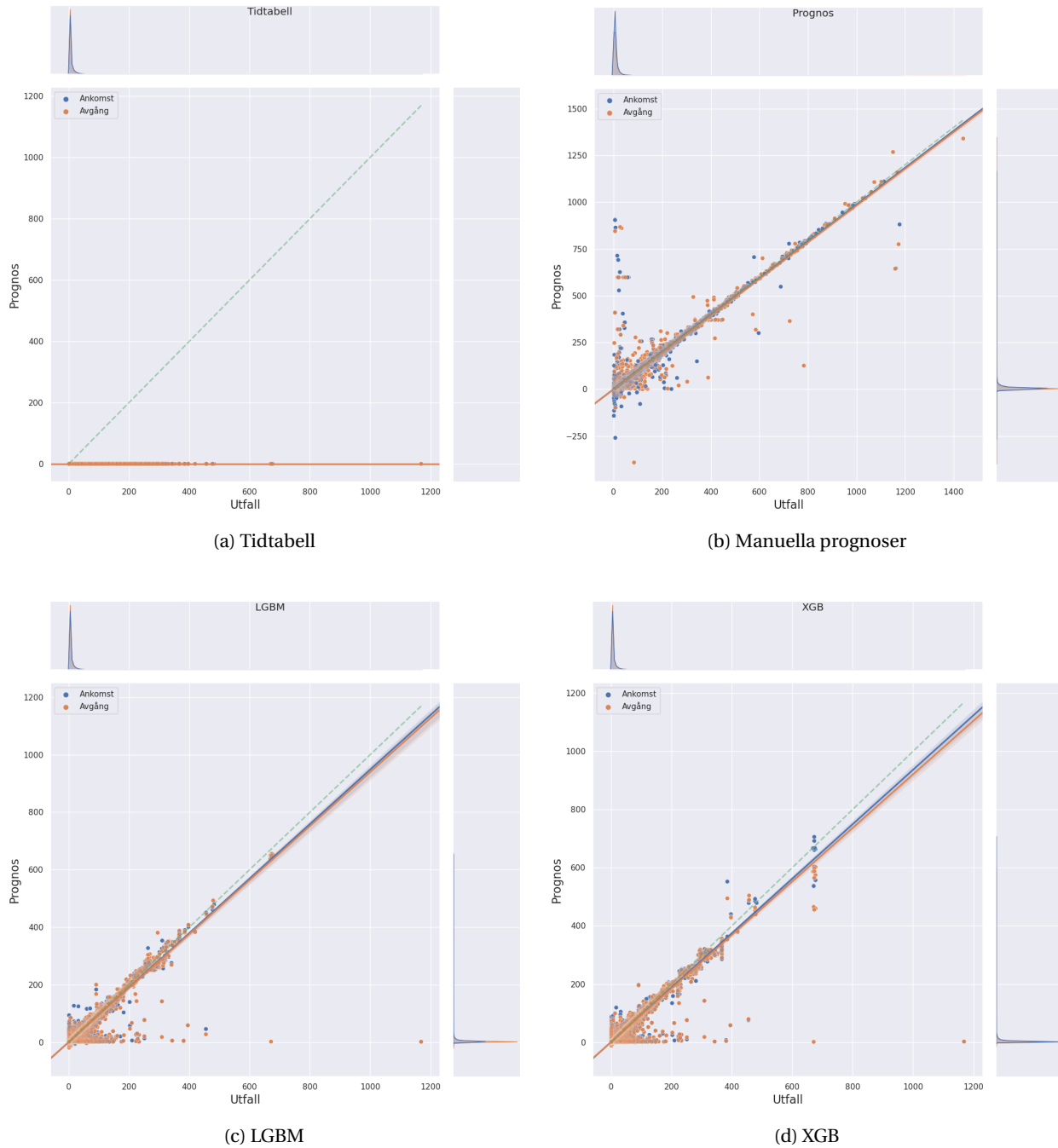
4.4.1 För Nästkommande Station

I Tabell 4.2 visas prestanda för ankomst -och avgångsprognoser för respektive modell och baselines. Prestanda är uttryckt i MAE (Mean Absolute Error), MSE (Mean Squared Error) och RMSE (Root Mean Squared Error). Vidare visas antal prognoser som genomförts under tidsperioden. Antal prognoser skiljer sig mellan manuella prognoser och resterande metoder då manuella prognoser inte alltid görs, samtidigt som de andra metoderna täcker varje ankomst och avgång över tidsperioden.

	Metod	MAE	MSE	RMSE	Antal
Ankomstförsening	Tidtabell	5.369 ± 13.498	211.026 ± 3472.142	14.527 ± 58.925	625729
	Manuell prognos	1.423 ± 6.611	45.732 ± 4015.78	6.763 ± 63.37	204280
	XGB	0.642 ± 2.892	8.774 ± 1769.285	2.962 ± 42.063	625729
	LGBM	0.817 ± 2.847	8.772 ± 1773.44	2.962 ± 42.112	625729
Avgångsförsening	Tidtabell	5.514 ± 13.569	214.522 ± 3522.283	14.647 ± 59.349	625729
	Manuell prognos	1.719 ± 6.544	45.776 ± 3599.486	6.766 ± 59.996	195444
	XGB	0.741 ± 3.21	10.855 ± 1864.271	3.295 ± 43.177	625729
	LGBM	0.896 ± 3.114	10.502 ± 1868.045	3.241 ± 43.221	625729

Tabell 4.2: Prestanda för samtliga metoder gällande nästkommande station uttryckt i MAE (Mean Absolute Error), MSE (Mean Squared Error) och RMSE (Root Mean Squared Error). Antal prognoser som genomförts för tidsperioden är också uttryckt. Manuella prognosers prestanda reflekterar den sista prognosen som görs innan ankomst/avgång till/från station.

I Figur 4.7 kan vi se hur prediktionerna förhåller sig till utfallet för nästkommande station utmed Sveriges tåg-
nät. Ett optimalt system skulle generera prognoser lika med utfall, vilket skulle betyda att samtliga punkter ligger
längs den gröna diagonalen. I Figur 4.7a ses titabellen som alltid predikterar att alla tåg har 0 minuters försening.
I Figur 4.7b ses hur manuella prognoser förhåller sig till dess utfall. Vad som kan urskiljas i denna figur är att man
i manuella prognoser ibland sätter negativa förseningsminuter. Detta är inget som uppkommer i resterande me-
toder då tidtabellen alltid predikterar 0 minuters försening och AI-modellerna (XGB och LGBM) är tränade på ett
dataset som endast innehåller positiva förseningsminuter, vilket också är sant för evalueringsdatasettet. I Figur
4.7c och Figur 4.7d syns respektive AI-modells prognoser för nästkommande station i relation till utfallet.

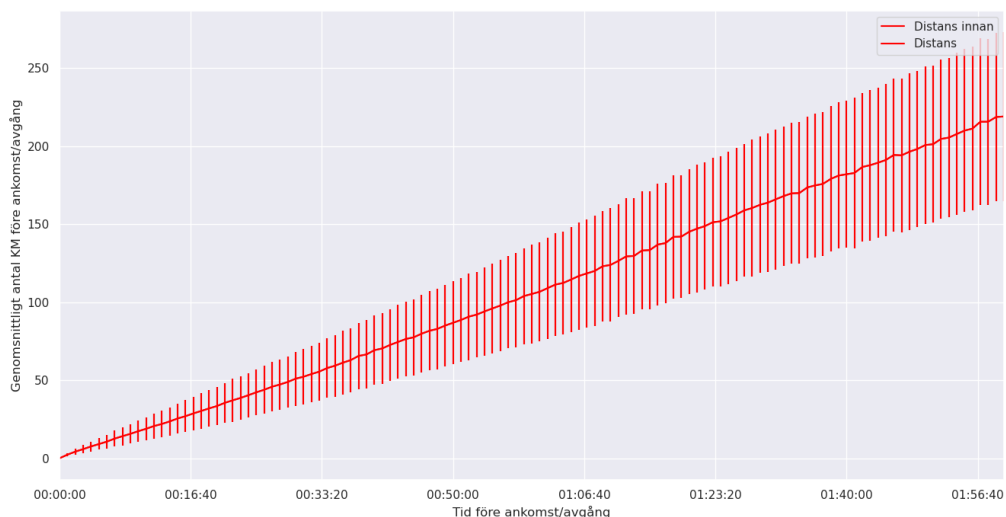


Figur 4.7: Relation mellan utfall och prognoser för tidtabellen, manuella prognoser, LGBM och XGB. Önskad relation är att utfall är lika med prognosen och att samtliga punkter ligger längs den gröna diagonalen. I samtliga plottar är ankomstförsening utmärkt med blå punkter och avgångsförseningar utmärkt med orange punkter.

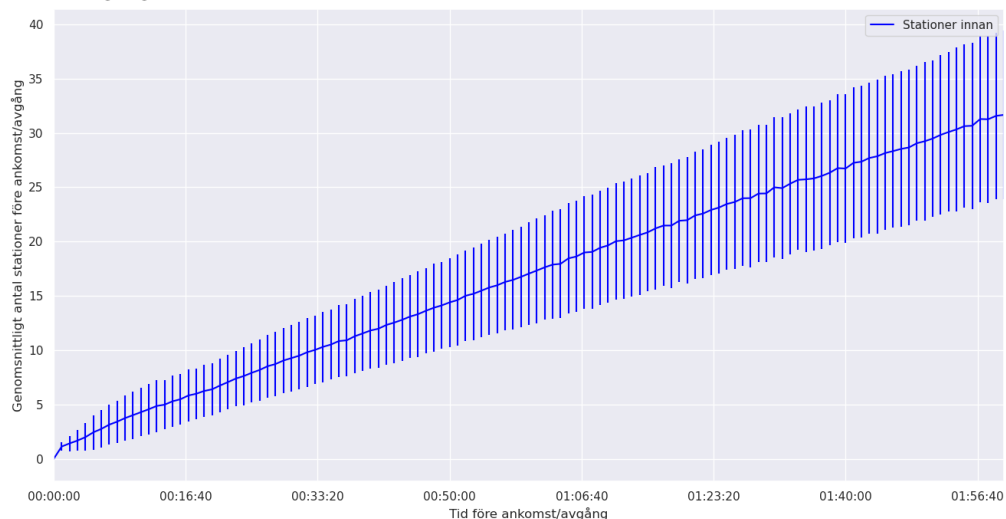
4.4.2 Framtida prognostering

Framtida prognoser i tågnätet kan beskrivas i tre olika dimensioner och enheter - tidsmässig framtid, distansmässig framtid och stationsmässig framtid. I Figur 4.8 visas hur tid innan ankomst/avgång förhåller sig till genomsnittlig distans innan ankomst/avgång respektive genomsnittligt antal stationer innan ankomst/avgång. Ur detta förhållande kan man utläsa att 2h innan ankomst/avgång så befinner sig tåget genomsnittligt ungefär 220km ifrån ankomst/avgångs-stationen. Man kan också utläsa att 2h innan ankomst/avgång har tåget i snitt 32 stationer kvar att passera innan ankomst/avgång.

I Figur 4.9 visas antalet prognoser som genomförts med hjälp av AI jämfört med antalet manuella prognoser som genomförts under samma tidsperiod. Antalet manuella prognoser som genomförts under evalueringsperioden är 1205434st och antalet AI-genererade prognoser som genomförts är 4560260st. Med det sagt, är AI-genererade prognoser skalbara.

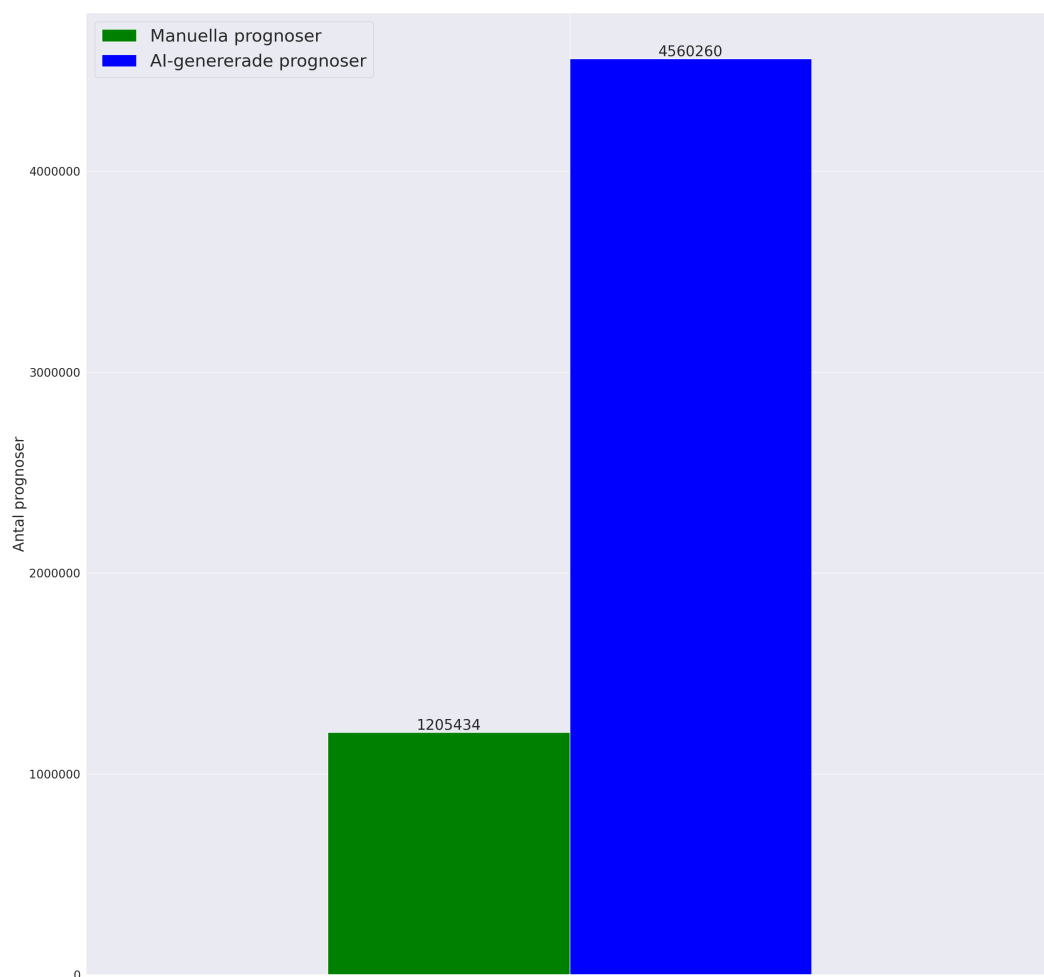


(a) Förhållande mellan tid innan ankomst/avgång och genomsnittlig distans (kilometer) innan ankomst/avgång.



(b) Förhållande mellan tid innan ankomst/avgång och genomsnittligt antal stationer innan ankomst/avgång.

Figur 4.8

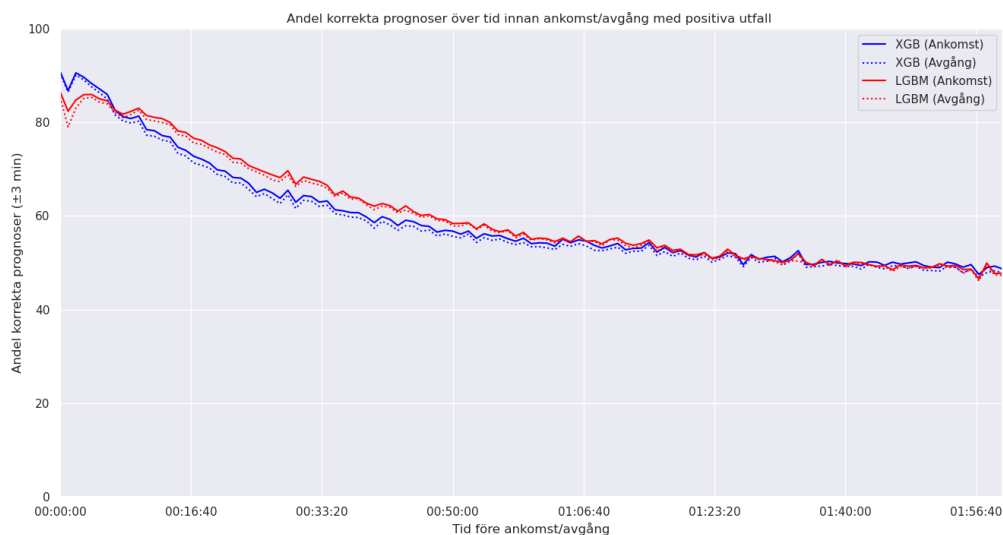


Figur 4.9: Antal prognoser som genomförts manuellt (grön stapel) och genom respektive metod för AI-genererade (blå stapel) över evalueringsperioden. Antal prognoser som gjorts enligt tidtabell är ej presenterat, men kan definieras som lika många som antalet AI-genererade prognoser.

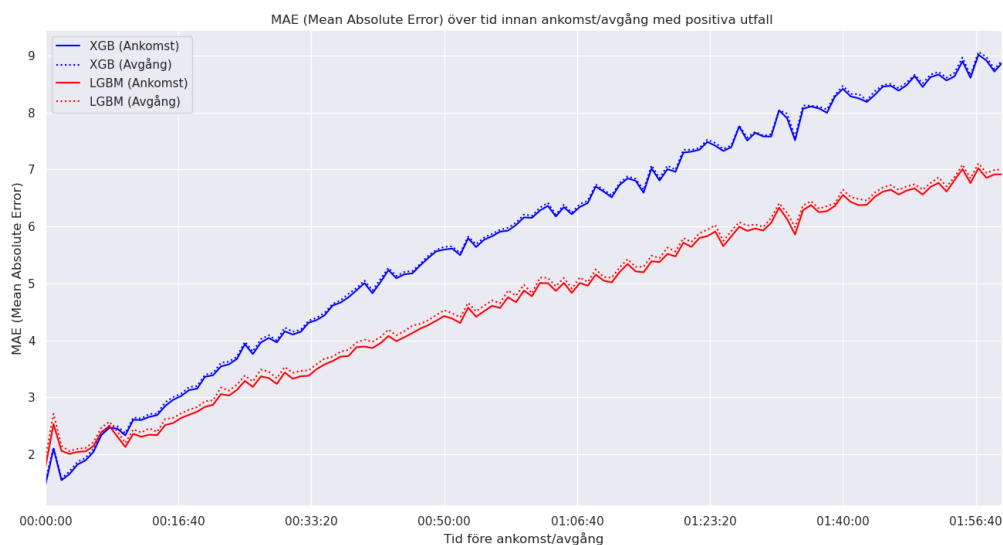
Tid

I Figur 4.10 presenteras resultaten för AI-modellerna (LGBM och XGB) gällande framtida prognoser som genomförts enligt metodik beskrivet i avsnitt 4.3.3 ur ett tidsbaserat synsätt. I graf 4.10a och 4.10b visas tid innan ankomst/avgång i x-axeln. I graferna 4.10a och 4.10b har tiden kvar till ankomst/avgång avrundats till närmsta minut. I graf 4.10a visas andel korrekta prognoser över tid innan ankomst/avgång och i graf 4.10b visas mean absolute error över tid innan ankomst/avgång. Korrekta prognoser har definierats som prognos ± 3 min innanför utfallet och andelen är angiven i procent, varpå ett högt värde så nära 100% som möjligt efterfrågas. Vidare, mean absolute error (MAE) eftersöks att vara så lågt som möjligt och har enheten minuter.

Ur graferna 4.10a och 4.10b kan man utläsa att vid två timmar innan ankomst/avgång så genererar AI-modellerna ca 50% korrekta prognoser och dessa prognoser har ett genomsnittligt fel på 7, respektive 9 minuter för LGBM och XGB. Andelen korrekta prognoser ökar och genomsnittligt fel minskar desto närmare ankomst/avgång tåget befinner sig vid prognossättning.



(a) Korrekt andel prognoser (± 3 min) över tid innan ankomst/avgång. Resultaten visar endast AI-modellerna.



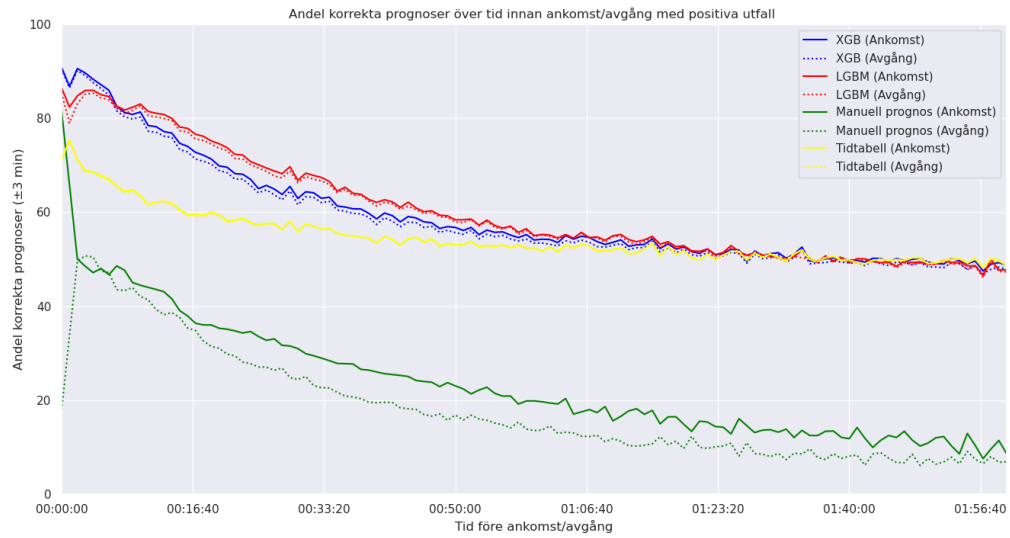
(b) Mean absolute error över tid innan ankomst/avgång. Resultaten visar endast AI-modellerna.

Figur 4.10: Röda linjer representerar prognoser genom LGBM och blåa linjer representerar prognoser genom XGB. Helsträckade linjer representerar ankomstprognoser och prickade linjer representerar avgångsprognoser.

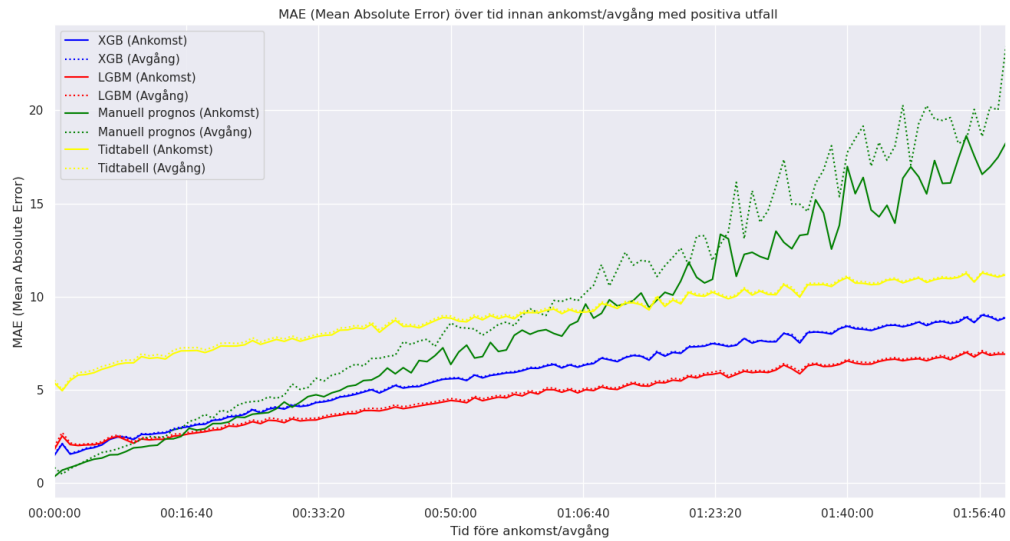
I Figur 4.11 visas, utöver AI-modellernas resultat, också tidtabellen och manuella prognosers resultat över tid innan ankomst/avgång. Tidtabellens prognoser har beräknats på samma sätt som AI-modellernas fast med ett output som alltid predikterar 0 minuters försening. Vidare har endast försenade tåg tagits med i evalueringssettet. Ur graf 4.11a och 4.11b kan man utläsa att manuella prognoser har, generellt, en lägre andel korrekta prognoser än de andra metoderna, men att andelen korrekta ankomstprognoser ökar drastiskt desto närmare ankomst tåget är. Man kan också se att andelen korrekta avgångsprognoser minskar drastiskt ju närmare avgång. Däremot, i Figur 4.11b kan man utläsa att det genomsnittliga felet vid prognoser inom 1h innan ankomst/avgång är desto lägre än tidtabellen och blir lägre än AI-modellernas vid ca 10 minuter innan ankomst/avgång.

Viktigt att notera i resultaten i graferna 4.11a och 4.11b är att antalet manuella prognoser som genomförts är betydligt färre än antalet prognoser som genomförts med andra metoder, vilket visas i Figur 4.9. För en mer rättvisare jämförelse bör man undersöka resultaten för respektive metod där manuella prognoser har genomförts, vilket presenteras i Figur 4.12.

I graf 4.12a och 4.12b visas resultaten endast för de ankomster och avgångar som också har en manuell prognos genomförd vid sig. Ur dessa grafer kan man utläsa att de manuella prognoserna överträffar tidtabellen på ett betydligt mer distinkt sätt än i graf 4.11a och 4.11b. I graf 4.12a och 4.12b ser vi också att AI-modellernas prestanda är bättre än i 4.11a och 4.11b men desto mer ostadig. Ostadigheten i AI-modellernas prestanda lär bero på den lägre andel prognoser som utforskas. Man ser också att diskrepansen i prestanda för ankomst och avgång ökar och att AI-modellerna har betydligt svårare att prediktera avgång för dessa fall.

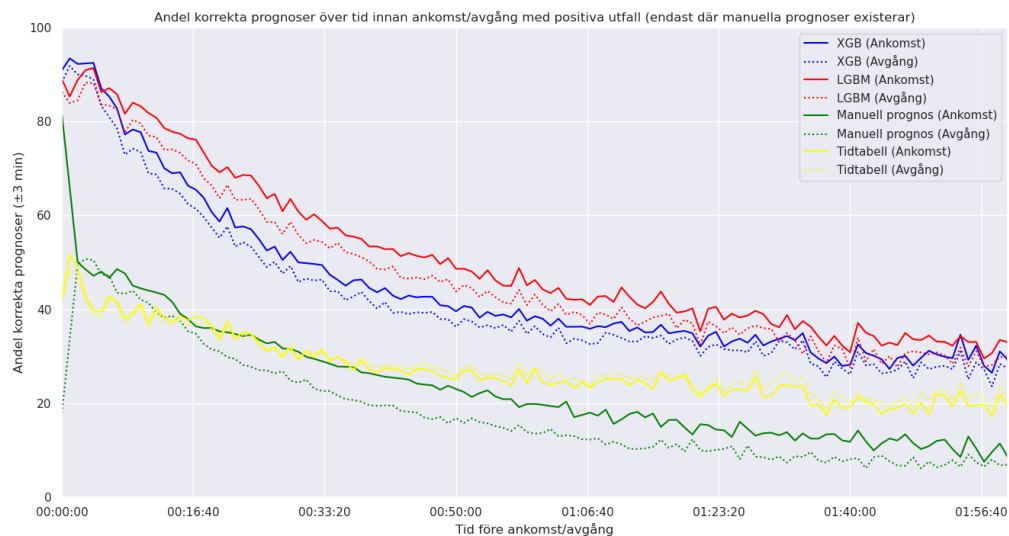


(a) Korrekt andel prognoser (± 3 min) över tid innan ankomst/avgång. Resultaten visar AI-modellerna och baselines.

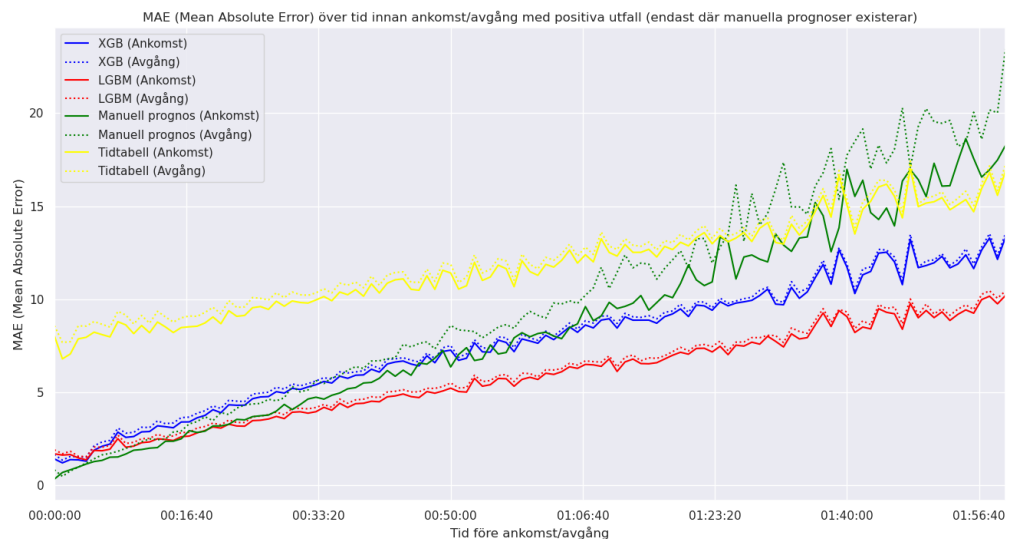


(b) Mean absolute error över tid innan ankomst/avgång. Resultaten visar AI-modellerna och baselines.

Figur 4.11: Röda linjer representerar prognoser genom LGBM och blåa linjer representerar prognoser genom XGB. Gröna linjer representerar manuella prognoser och gula linjer representerar tidtabellen. Helsträckade linjer representerar ankomstprognoser och prickade linjer representerar avgångsprognoser.



(a) Korrekt andel prognoser (± 3 min) över tid innan ankomst/avgång. Resultaten visar AI-modellerna och baselines.



(b) Mean absolute error över tid innan ankomst/avgång. Resultaten visar AI-modellerna och baselines.

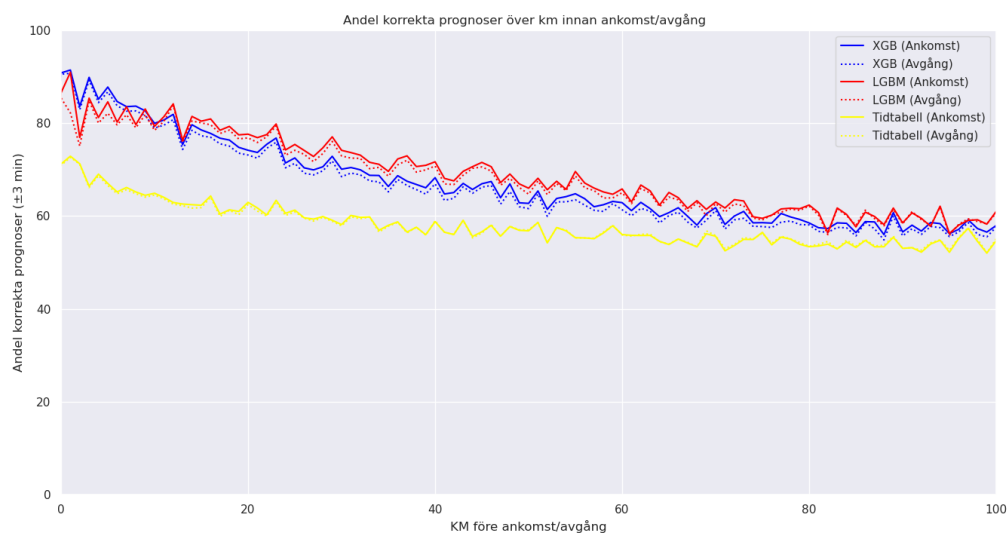
Figur 4.12: Röda linjer representerar prognoser genom LGBM och blåa linjer representerar prognoser genom XGB. Gröna linjer representerar manuella prognoser och gula linjer representerar tidtabellen. Helsträckade linjer representerar ankomstprognoser och prickade linjer representerar avgångsprognoser.

Distans

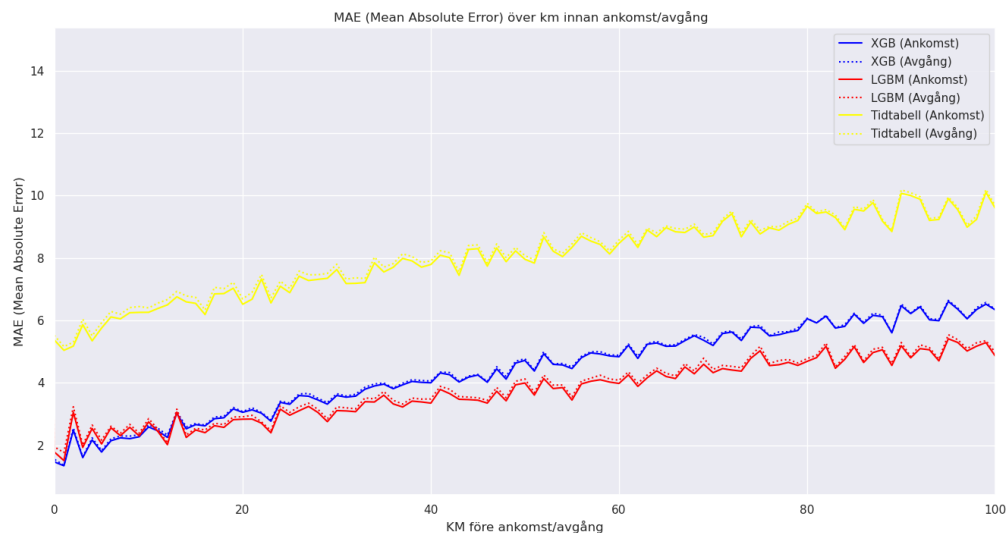
I Figur 4.13 visas prestanda för AI-modellerna XGB och LGBM samt tidtabellen över kilometer innan ankomst/avgång. I graf 4.13a visas andel korrekta prognoser över kilometer innan ankomst/avgång och i graf 4.13b visas mean absolute error över kilometer innan ankomst/avgång.

I dessa grafer är, till skillnad från graferna i Figur 4.10 - 4.12, manuella prognoser inte med. Detta för att det dataset som projektet tillhandahållits, inte beskriver vart någonstans de manuella prognoser har utförts, utan endast när dem har utförts. Därför kan man inte på ett säkert sätt beskriva hur långt kvar tåget har vid utförd manuell prognos, varpå den enda baseline som går att använda är tidtabellen.

I graferna 4.13a och 4.13b har avståndet kvar till ankomst/avgång avrundats till närmsta kilometer. Graf 4.13a och 4.13b har en liknande karaktäristik som tidigare grafer och visar på att AI-modellerna har en generellt sätt högre andel korrekta prognoser och ett generellt sätt lägre genomsnittligt fel än tidtabellen.



(a) Korrekt andel prognoser (± 3 min) över distans innan ankomst/avgång. Resultaten visar AI-modellerna och tidtabell.

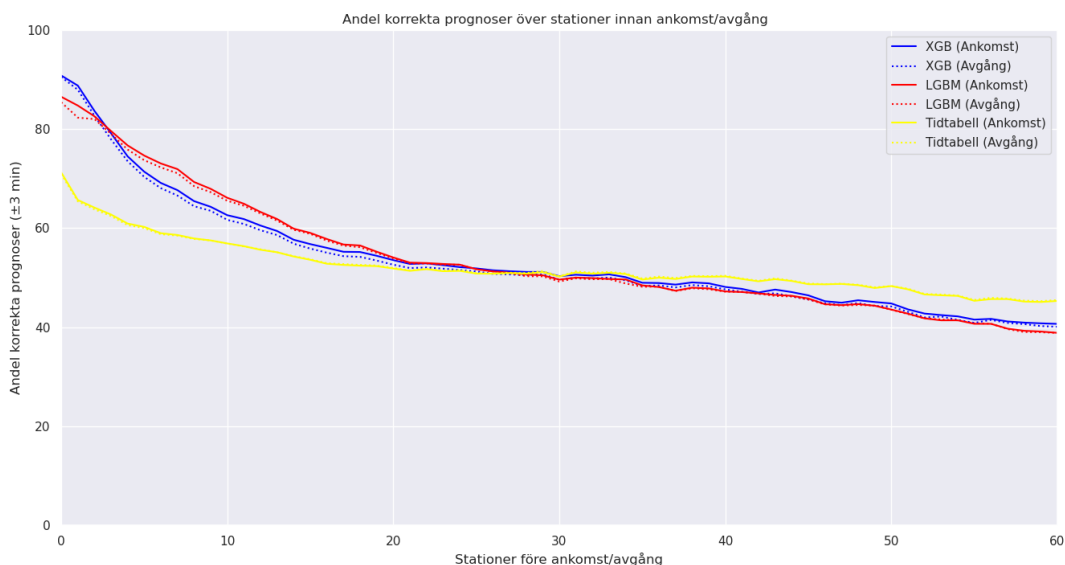


(b) Mean absolute error över distans innan ankomst/avgång. Resultaten visar AI-modellerna och tidtabell.

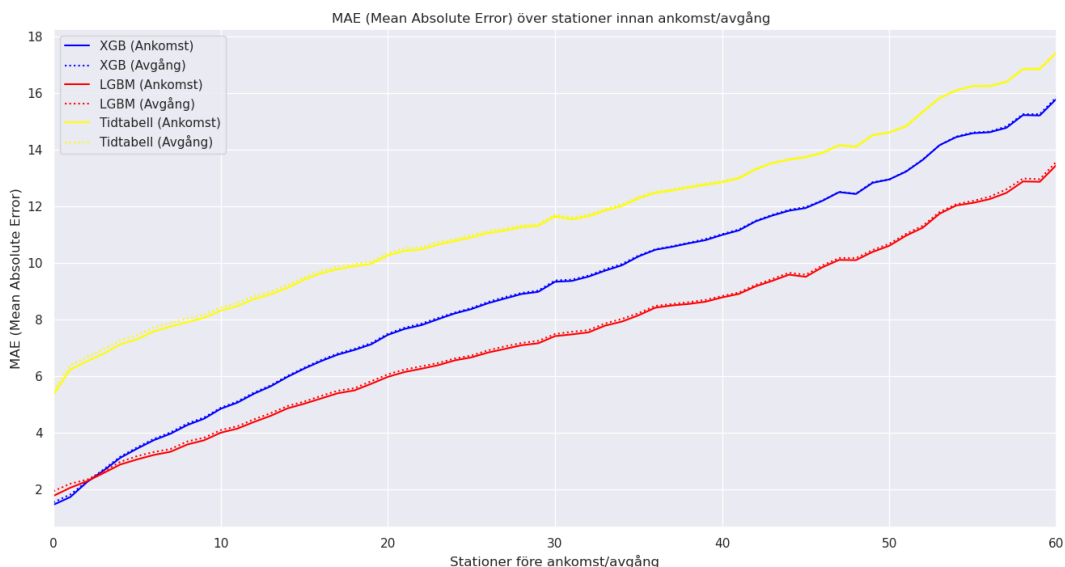
Figur 4.13: Röda linjer representerar prognoser genom LGBM och blåa linjer representerar prognoser genom XGB. Gröna linjer representerar manuella prognoser och gula linjer representerar tidtabellen. Helsträckade linjer representerar ankomstprognoser och prickade linjer representerar avgångsprognoser.

Stationer

I Figur 4.14 visas prestanda för AI-modellerna XGB och LGBM samt tidtabellen över stationer innan ankomst/avgång. Antal stationer innan ankomst definieras som hur många stationer det finns mellan plats vid prognos och ankomst- och avgångsplats. Därav går x-axeln ner till 0 stationer då detta representerar prognoser som sätts vid stationen innan ankomst/avgång. Graferna 4.14a och 4.14b har liknande karakteristik som tidigare grafer men påvisar en alltmer jämnare kurva. Detta för att granulariteten (stationer) är lägre än vid tidigare jämförelser (minut och kilometer).



(a) Korrekt andel prognoser (± 3 min) över stationer innan ankomst/avgång. Resultaten visar AI-modellerna och tidtabell.



(b) Mean absolute error över stationer innan ankomst/avgång. Resultaten visar AI-modellerna och tidtabell.

Figur 4.14: Röda linjer representerar prognoser genom LGBM och blåa linjer representerar prognoser genom XGB. Gröna linjer representerar manuella prognoser och gula linjer representerar tidtabellen. Helsträckade linjer representerar ankomstprognoser och prickade linjer representerar avgångsprognoser.

Kapitel 5

Vädrets påverkan på järnvägssystemet

För att utforska hur vädret påverkar järnvägen behöver forskningsfrågan delvis omformuleras. Begreppet "väder" är brett och måste definieras närmare för att vara mätbart. Kunskap om vanliga meteorologiska variabler är förankrad hos såväl professionella användare som den bredare kretsen av användare av väderinformation. I detta forskningsprojekt har närliggande domäner använts för att komma närmare målet som kan benämnas "järnvägsväder". "Järnvägsväder" är ett samlingsnamn för de samband man eftersträvar att identifiera med fokus på och anpassning till ändamålet - alltså järnvägen - snarare än ren meteorologisk information. Denna anpassning till domänen "järnväg" fordrar introduktion av nya variabler som närmare kan beskriva följderna av vädret snarare än tillståndet i atmosfären, d.v.s. "väder". Forskningsfrågan behöver därför omformuleras och utökas från att gälla "väder" till: Vilken kombination av meteorologiska variabler och andra variabler, inklusive värden på dessa variabler, har påverkan på järnvägssystemet, och hur ser sambanden ut? Vidare måste vi även definiera vad som åsyftas med "påverkan", vilket kan göras genom introduktion av en rad tillgängliga variabler som på olika sätt beskriver påverkan. Exempel på sådana variabler är "trafikpåverkans längd" och "merförsening".

Jämförelser mellan vädervariabler och variabler associerade med händelser i järnvägsnätet kräver en strategi för att matcha väderdata mot händelsedata i tid och rum. Tidsaspekten och det geospatiala beroendet ger därför upphov till nya frågor som rör jämförelsen mellan tidpunkten för inträffade händelser och väderdatats fixa tidsintervall, samt den geospatiala presentationen av väder- respektive händelsedata. Medan tidsaspekten är en förhållandevis enkel definitionsfråga, kräver den geospatiala anpassningen en mer omfattande anpassning. Denna beskrivs närmare under nästkommande avsnitt om metoder, en kartläggning av delsträckor i järnvägsnätet anpassad till väderdatat och forskningsprojektets behov.

5.1 Väderdata

Väderdata kan vara observationer eller prognoser. Medan observationer förekommer på specifika platser genom mätningar av ett begränsat antal meteorologiska variabler, är väderprognoser resultat från olika numeriska modeller som körs med flera observationskällor som indata, numerisk behandling och numeriska metoder som tillämpas på fysikaliska ekvationer där tid ingår som nyckelvariabel. Outputen från en numerisk väderprognosmodell utgörs av beräknade värden för ett relativt stort antal meteorologiska variabler, för ett antal tidssteg fram i tiden, för ett stort antal punkter tillhörande ett gridnät som spänner över ett geografiskt område.

Väderdata som användes i projektet utgjordes av historiska prognoser från den numeriska väderprognosmodellen GFS från NOAA ¹, och därutöver algoritmer för beräkning av nya variabler som beskriver förlopp på marken, framtagna av ShareWeather ². Meteorologiska data hanterades i systemet ShareWeather, ett Javabaserat system som m.h.a. mjukvaran netCDF hanterar modelloutput i formatet GRIB, kan utföra nya beräkningar i Java och genom en webbserver serva bl.a. webbaserade analysverktyg som användes i projektet. Flödet kan illustreras genom detta exempel: Meteorologiska variabler (1) som nederbörd, temperatur, avdunstning, molnmängd, används för att beräkna nya variabler (2) som nederbördsmängd i form av regn, snöblandat regn respektive snö, mängden nysnö, snö som vid vissa vindhastigheter driver in från omkringliggande terräng, mängden nysnö som smälter på

¹<https://www.ncei.noaa.gov/products/weather-climate-models/global-forecast>

²<https://shareweather.com>

marken till vatten, mängden vatten på marken orsakad av eventuell liggande snö i terräng som smälter, slutligen mängden vatten på marken som fryser till is som resultat av alla de processer som ägt rum under de föregående timmarna inklusive avdunstning. Genom detta exempel illustreras även fördelarna med att använda modelloutput i kombination med nya beräkningar, vilket kan jämföras med observationsdata som tillför ett, i relation till modelldata och beräkningar, begränsat antal variabler. Om man istället hade använt observationsdata från VViS hade tillgängliga data i exemplet ovan varit: total nederbördsmängd, nederbördsslag, beläggning. I det senare fallet ställs man även inför utmaningen att utföra en geospatial anpassning mellan vägnätet och järnvägsnätet, vilket ställer nya krav på hantering av höjdskillnader och andra lokala skillnader. Givetvis krävs geospatial anpassning även mellan modellpunkter och platser med inträffade händelser (se Figur 3.2), men med modelldata kan flera nya beräkningar göras för att utföra en anpassning som i bästa möjliga mån motsvarar platser för inträffade händelser i järnvägsnätet. Vidare kan man göra prognoser av framtida tillstånd, medan observationsdata endast beskriver nuläget.

5.1.1 Geospatial anpassning och kartläggning av delsträckor i järnvägsnätet

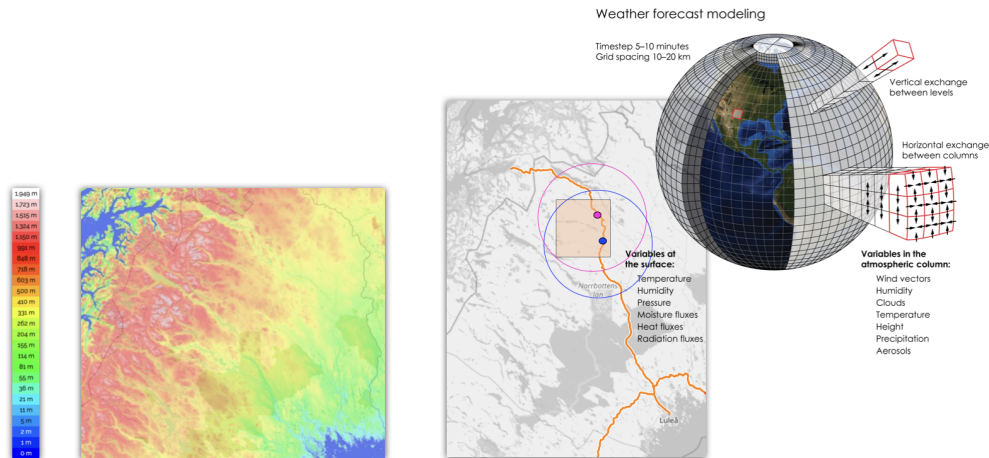
Sveriges järnväg sträcker sig 15 600 km. I detta vidsträckta nät inträffar flera händelser i söder än i norr. En del järnvägssträckor har ringa eller inga nämnvärda höjdskillnader, medan det förekommer stora höjdskillnader på en del sträckor. Det globala vädret karakteriseras av att luftmassor med liknande meteorologiska egenskaper täcker stora geografiska områden. I gränzonen mellan luftmassor skapas vädersystem och mer dynamiska väderförhållanden med väderomslag som kan ske relativt snabbt (under timmar snarare än dygn). Givetvis kan flera vädersystem avlösa varandra och på så sätt bidra till dagliga dynamiska förändringar.

I en luftmassa är höjd över hav avgörande för temperaturens värde. Det gäller såväl marknära nivå (exempelvis 2 meters höjd som i det synoptiska observationsnätet upprättat enligt meteorologisk standard) som marktemperatur (i VViS: yttemperatur). I det senare fallet, ytnivå, kan även den omgivande miljön, liksom andra processer kopplade till markegenskaper, spela in. Eftersom modellen har en spatial upplösning på upp till 20 km kommer dess output att ge värden som beskriver ett genomsnittligt värde för en hel gridruta (20 km x 20 km). Den järnvägssträcka som löper igenom en ruta kommer dock i de flesta fall att variera i höjd, vilket resulterar i olika temperatur beroende av aktuell höjd över hav.

Temperaturkorrigering med avseende på höjd över hav består av en enkel operation och används exempelvis inom flygtransport. Det gäller samtidigt att, även om korrektionen är enkel att utföra, är det sällan atmosfären exakt uppfyller de villkor man utgår ifrån vid denna korrigering. I själva verket kan, beroende av vädersituation, de verkliga förhållanden som råder i atmosfären ha en högst varierande temperaturhöjdprofil. För att uppnå en rimlig noggrannhet under en grad ($< 1\text{ }^{\circ}\text{C}$), helst tiondels grader, är denna korrigeringsmetod dock ganska tillfredsställande. Man kan utgå ifrån att temperaturavtagandet i en neutralt skiktad atmosfär är ca $0.7\text{ }^{\circ}\text{C}$ per 100 m. I den korrigering som använts här har hänsyn tagits även till ytterligare några faktorer kopplade till vädertyp (exempelvis molnighet, luftfuktighet och atmosfärens stabilitet), vilket gör att temperaturkorrigeringen varierade mellan ca 0.5 och $1.0\text{ }^{\circ}\text{C}$ per 100 m höjd. Detta beror på att det system som använts för att få en delvis anpassad väderprognos (ShareWeather) använt något mer avancerade algoritmer som tar hänsyn till flera faktorer. Återigen, med hänvisning till den noggrannhet som är rimlig att eftersträva, landar man i en skillnad som ligger inom tiondels grader, förutsatt att höjddatat är adekvat.

Temperaturen påverkar många andra förlopp. Exempelvis avgör lufttemperaturen nederbördsslag, som andelen regn respektive snö. Yttemperaturen bestämmer sedan exempelvis om nederbörden smälter (plusgrader på marken vid snöfall) respektive fryser på (regnvatten som i ett senare skede kan frysa till is om yttemperaturen faller under noll), eller om ackumulerad snö i terrängen börjar smälta. Skiftande temperaturer med höjden påverkar även luftfuktigheten och risken för markbeläggningar vilket man exempelvis kan erfara när vattendroppar kondenserar respektive sublimerar och bildar moln, daggdroppar eller frost, vid sänkt temperatur. Utöver temperaturkorrigering är det därför önskvärt att utföra en rad andra omberäkningar, för att återge en så realistisk bild som möjligt av de lokala skeendena på en plats som befinner sig på en högre respektive lägre höjd än den position för vilken vi har en observation eller prognos. Medan detta kan visa sig vara svårt om man exempelvis har ett stort antal datapunkter, som i en AI-baserad modell, har det i fokusområde väder inte funnits hinder att utföra nya beräkningar, särskilt som det system som använts (ShareWeather) redan har algoritmer för att utföra beräkningar av denna typ.

Delsträckor valdes utifrån en kartläggning av järnvägsnätet som utgick ifrån de geografiska och topografiska förutsättningarna. För varje delsträcka valdes en representativ höjd över hav. På så sätt kunde ca hundra delsträck-



Figur 5.1: Väderprognosmodellens gridnät anpassades till delsträckor med homogena egenskaper som definierades för projektet, bl.a. i syfte att utföra anpassning m.a.p. höjd över hav.

kor identifieras. Figur 5.1 illustrerar principen för hur delsträckor, som definierats som punkter med en position och höjd över hav, matchas mot en ruta i prognosmodellens gridnät. När en delsträcka tilldelas prognosvärden hörande till en viss prognosgridruta utförs korrigeringar baserade på höjdskillnaden mellan delsträcka och gridruta, varvid omberäkningar av flera variabler sker.

Denna anpassning kan ses som ett första steg i anpassningen av väderprognoser till järnvägsnätet. I själva verket är det önskvärt att i det framtida arbetet med "järnvägsväder" införa flera anpassningar relaterade till den närliggande terrängen, som exempelvis påverkas av öppna vattenytor och terrängprofil.

5.2 Metoder och vetenskaplig metodik

5.2.1 Metodologiska utmaningar och vetenskaplig metodik

Det finns en grundläggande utmaning när man studerar väder och det är att vädervariablerna inte går att kontrollera. Samtliga variabler varierar samtidigt, medan vi i stället eftersträvar att variera endast en variabel i taget. Verklighetens vädersituationer erbjuder inte ideala förhållanden för att sätta upp vetenskapliga studier. Exempelvis kommer temperaturen ofrånkomligt att variera i bakgrunden medan vi studerar sambandet mellan vind och händelser i järnvägssystemet. Konkret kan detta, såvida man inte är försiktig när man tolkar resultaten, leda till felaktiga tolkningar och slutsatser. I detta exempel är vindstyrka en variabel som är beroende av tryck (eller tryckfördelningar över ett område). Samtidigt är det känt att situationer med låga temperaturer ofta uppträder vid tämligen stationära väderlägen nära högtryckscentrum och svaga vindar som följd. Detta ger förhållandet svaga vindar vid kyla. Men om både kyla (låga temperaturer) och kraftig blåst (höga vindhastigheter) påverkar järnvägssystemet kommer vi kanske att felaktigt dra slutsatsen att situationer med svag vind leder till större påverkan, när det i själva verket är en annan variabel, kyla, som kanske orsakat påverkan. Denna typ av metodologiska utmaning var konstant gällande under projektet. Problemet illustreras i Figur 5.2 a), varav det framgår att låga vindhastigheter kan inträffa vid kyla, vilket var vanligt förekommande i norra Sverige under januari 2021. Samtidigt fanns i mellersta Sverige flera tillfällen med svaga vindar och nederbörd.

Den andra utmaningen är själva förekomsten av vissa väderfenomen. Exempelvis är stormvindar relativt sällsynta och temperaturer rör sig inom rimliga intervaller för en viss årstid för ett visst geografiskt område. Att vissa väderfenomen är mindre, och andra mer, vanliga resulterar i få datapunkter för vissa värden av vädervariabler och kombinationer av variabler och flera datapunkter för andra vädervariabler. I en kvantitativ studie erfordras upp mot hundra mätvärden eller strax därunder. Datamängder med få mätpunkter kan visserligen indikera mönster, men med exempelvis ett trettiotal mätningar att analysera bör man istället använda kvalitativa metoder.

Vidare uppmanade studierna allteftersom de gav flera resultat, till en allt mer granulär uppdelning både geo-



Figur 5.2: a) Bilden visar avvikelser från medelmerförseningen (x-axeln) för några enkla meteorologiska variabler, där gult är vindhastigheter, blått nederbörd, lila kyla (d.v.s. låga temperaturer). Frost, som är en beräknad variabel, illustreras i vitt. När merförseningen är större än medel (d.v.s. när påverkan är större) är värdet positivt, i annat fall negativt. Y-axeln representerar olika latituder, d.v.s. låga värden på y-axeln motsvarar södra Sverige medan höga värden representerar norra delarna av landet. b) Illustration av hur studien var uppbyggd i sin helhet, där kvalitativ och kvantitativ analys användes omväxlande i en iterativ process.

grafiskt (driftledningsområden) och på orsakskodnivå. Detta resulterade i sin tur i färre datapunkter för varje uppdelning, vilket ledde till viss försiktighet vid hantering och tolkning av resultaten.

Å andra sidan ger flera liknande studier av samma fenomen tillfälle att jämföra erhållna resultat från olika studier med varandra. Är resultaten konsistenta pekar det även på en viss tillförlitlighet. I de studier som gjordes analyserades exempelvis en månad var för sig innan samtliga månader slogs ihop till ett resultat för en hel tidsperiod på flera månader. Vissa resultat kunde analyseras både kvantitativt och kvalitativt för att kontrollera intressanta eller avvikande värden. Detta arbete resulterade inte helt sällan i att en ny vädertyp kunde definieras och läggas till, eller att existerande vädertyper modifierades och/delades upp i flera vädertyper.

5.2.2 Kvalitativa och kvantitativa metoder

I en kvantitativ studie är data strukturerade och kan kvantifieras i siffror eller kategorier. Den senare typen av data, kategoriska data, användes för att definiera vädertyper, medan händesedata var representerade i siffervärden. I de flesta delstudierna analyserades orsakssamband mellan en väderrelaterad kategori, eller "vädertyp", och variabler uttryckta i siffervärden som exempelvis trafikpåverkans längd och merförsening (båda i antal minuter), eller händelsefrekvens (exempelvis antal händelser per dygn). Har man få mätvärden kommer resultatet att vara behäftat med betydligt större osäkerhet.

Kvalitativa metoder innebär att man studerar och läser (en uppsättning av) data i detalj för att på så sätt hitta mönster och definiera samband, som man sedan testar på ett större dataset. För att initialt kunna identifiera vilka variabler som kan vara av intresse samlar man först på sig kunskap som redan pekar på vissa hypoteser om orsakssamband. Exempelvis fanns inom Trafikverket redan kunskap om att sträng kyla respektive kraftig vind kan leda till påverkan på järnvägssystemet. Under projektet utarbetades flera s.k. "vädertyper" som definierades som kombinationer av värden av flera variabler som inträffade samtidigt. Några vädertyper kunde således definieras redan i projektets inledningsfas. Flera nya vädertyper definierades med inspiration från den närliggande domänen "vägväder". Andra vädertyper var resultat av kvalitativ analys följt av kvantitativ analys av samband mellan en ny vädertyp och dess påverkan på järnvägssystemet. Tidsperioden i de inledande studierna var ofta en månad, som sedan utvidgades till flera månader. Detta gav också tillfälle att först analysera data kvalitativt för att definiera vädertyper och senare tillämpa dessa för att testa samband på ett större dataset. Ibland gav resultaten fingervisning om att nya vädertyper bör läggas till.

Kortfattat kan sägas att de kvalitativa metoderna i detta projekt utformade hypoteser om samband mellan väder och händelser i järnvägssystemet, medan de efterföljande kvantitativa metoderna användes för att få fram siffervärden som beskriver sambandet och samtidigt testar och eventuellt bekräftar hypoteserna. Studierna bedrevs i en iterativ process som genererade nya hypoteser (steg 1) och vädertyper för att därefter testa dem (steg 2), som illustrerat i Figur 5.2 b). Även om kvantitativa metoder ibland tillämpades även för att identifiera väder-typer kan man säga att det första steget i den iterativa processen oftast utgjordes av kvalitativa metoder, medan det andra steget oftast byggde på kvantitativa metoder, ibland dock kompletterade av kvalitativ analys när antalet



Figur 5.3: Analysverktyg som på samma tidslinje visar väder och händelser.

mätpunkter inte levde upp till kraven.

5.2.3 Analysverktyg

Efter anpassningar och införande av historiska prognoser i ett Javabaserat system med netCDF från ShareWeather³ servern, kunde en webbapplikation presentera både väderprognoser från väderprognossystemet ShareWeather och händelser i järnvägssystemet. Webbverktyget användes, utöver kvalitativ analys, för filtrering av data, olika beräkningar, visualisering och sammanställning av resultat.

I Figur 5.3 visas trafikpåverkans längd representerad på en karta i form av cirklar där radien visar påverkans storlek medan olika vädertyper är representerade m.h.a. olika färger (exempelvis blått för nederbörd, gult för vind, grönt för bra" väder). Analysverktyget i 5.3 visualiserar aktuellt väder också på en tidslinje och per tidssteg. Tidslinjen presenterar även händelser (röda staplar) och påverkan (staplar med svag opacitet för att man ska kunna följa påverkan från flera pågående händelser. I diagrammet, som brukar kallas meteorgram, finns även en realistisk återgivning av vädrets konsekvenser på markytan, som exempelvis snödjup i terräng, vattenansamling på marken, is- respektive frostbeläggningar. Värden, som exempelvis snödjup, återges genom skalan på y-axeln (negativa värden för mark). I meteorgrammets övre del visualiseras, utöver temperaturkurvor (yttemperatur, daggpunkt), även vindstyrkor och -riktning, molnighet, luftfuktighet och förekommande nederbördsslag. Tidslinjen visar vädret för en specifik delsträcka/punkt medan filtreringsfunktionerna möjliggör presentation av samtliga händelser inom valfritt avstånd från delsträckan/punkten, filtrering på text, orsakskod och även vädertyp närmast händelsen.

5.3 Variabler associerade med händelser

Liksom för hela AIRT projektet hämtades händelsedata i huvudsak från LUPP, där händelser som inte haft trafikpåverkan filtrerats bort. Flera variabler associerade med händelser studerades i början av projektet, varefter fokus kom att hamna på trafikpåverkans längd och framför allt händelsefrekvens. I de första studierna användes ett större dataset, närmare bestämt ett komplett dataset från LUPP, dock ingick ofta en kortare tidsperiod i studien, i de flesta fall en eller ett par månader, p.g.a. datamängden.

Tidsperioden omfattade som mest drygt fyra månader, från 20201023 till 20210228. Dataanalysen utfördes månadsvis vilket dels skapade förutsättningar för kvalitativ analys, dels möjliggjorde jämförelser mellan studierna för att säkerställa en viss tillförlitlighet.

Inledningsvis gjordes en förstudie av händelsedata för att få grepp om datat, varvid fokus inledningsvis lades på orsakskoder som antogs ha en nära koppling till väder. Relativt snart kom studierna att omfatta samtliga orsakskoder som studerades av övriga studierna i AIRT projektet, d.v.s. infrafel och fordonsfel.

Under projektet studerades först variabler som merförsening och händelsens varaktighet, varav den senare var definierad som förfluten tid mellan tidpunkten då händelsen registrerades och tidpunkten då händelsen avsluta-

³<https://shareweather.com>

des. Inledningsvis användes alla händelser som registrerats i LUPP, vilka, till skillnad från det filtrerade datat som användes i AI-modellen, bestod av samtliga händelser oavsett trafikpåverkan. Händelsens varaktighet visade sig dock inte vara fullt representativ och lades åt sidan till fördel för variablerna ”merförsening” och ”trafikpåverkans längd”. Under andra halvan av projektet flyttades fokus delvis mot händelsefrekvens, som ansågs vara den variabel som ur intressenternas synvinkel uppvisat intressanta samband.

Med hjälp av det analysverktyg som användes kunde olika filtreringar göras. Utöver exempelvis orsakskod, vädertyp och sökord i text, kunde man välja alla händelser inom ett avstånd från en vald punkt i järnvägsnätet, för vilken det även fanns en väderprognos. Detta gjorde det möjligt att laborera med avstånd och inkludera flera händelser, för att studera eventuell påverkan på järnvägsnätet orsakat av väderförhållanden som inträffat längre bort. Avstånd som testades var 5 km upp till 200 km. För att visualisera och studera alla händelser för tidsperioden samtidigt användes kartpresentation som bland annat kunde filtreras med avseende på vädertyp (närmast händelsen), text, orsakskod (5.3).

5.3.1 Merförsening

Variabeln merförsening studerades för december 2020 respektive januari 2021 och byggde på samtliga data från LUPP för samma period. En fördel här, i relation till de övriga studierna, är att man kunde studera händelser som filtrerats bort i det dataunderlag som användes i projektets övriga studier.

5.3.2 Trafikpåverkans längd

Trafikpåverkans längd var den variabel som studerades vid fokus på Prediktion av tidsåtgång för trafikstörande händelser, där AI-metoder tillämpades, och var därför av särskilt intresse för jämförelse mellan AI-modellen och de analyser som gjordes under studierna av samband mellan väder och händelser i järnvägssystemet. Samband mellan vädervariabler och trafikpåverkans längd studerades för perioden 20201023 - 20210228.

5.3.3 Händelsefrekvens

Även händelsefrekvens studerades för tidsperioden från 20201023 till och med 20210228. Utöver detta studerades händelsefrekvens inledningsvis för några utvalda orsakskoder över ett helt år, nämligen 2020.

5.3.4 Text

Vid Trafikverket skapas en manuell framställd text som beskriver inträffade händelser. Dessa texter analyserades kvalitativt. Interntexten användes i flera studier och för olika tidsperioder. Detta utgjorde ett komplement till de övriga metoderna både för framtagande av vädertyper och under analysarbetet.

5.4 Meteorologiska variabler

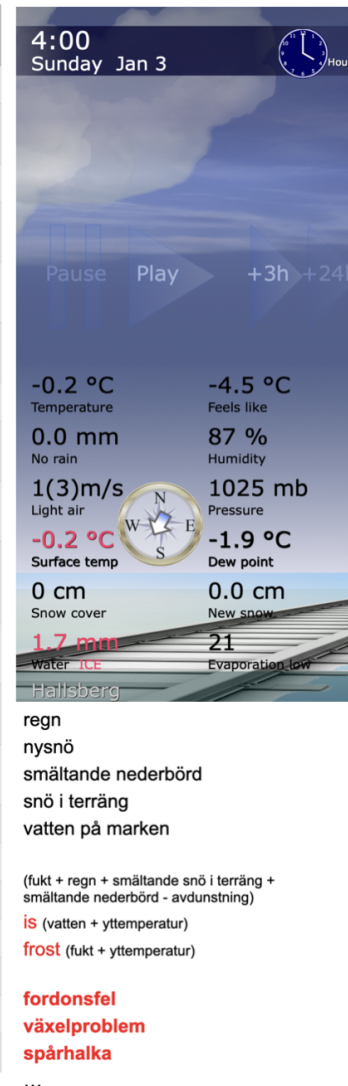
Målet med fokusområde Väder i detta projekt var ju att upptäcka och formulera samband mellan väder och händelser i järnvägssystemet, varför fokus riktades mot meteorologiska variabler, viktiga siffervärden, intervaller och eventuella tröskelvärden, kombinationer av flera variabler och värden av dessa.

5.4.1 Järnvägsväder

Utgångspunkterna i projektet var dels branschkunskap inom järnväg, dels tidigare forskning och kunskap avseende meteorologiska variabler och närliggande fält. För att beräkna konsekvenserna av vädret på marken eller rälsen räcker inte endast meteorologiska variabler eller meteorologisk kunskap. Kunskapen måste vidareutvecklas för att formulera samband anpassade till järnvägens behov.

En första anpassning utgörs av den spatiala anpassning av data som beskrevs tidigare, illustrerad i Figur 5.1. Temperaturen (och yttemperaturen) omberäknades med avseende på höjd över hav. Eftersom temperaturen påverkar andra förlopp gjordes i samband med höjdkorrigerad av temperatur även omberäkningar av en rad andra variabler som molnighet, relativ fuktighet, nederbördsslag, nederbördsmängd, andel nederbörd som faller som snö respektive regn, mängden nysnö som smälter direkt på marken (vid yttemperaturer större än noll), nysnö som

Meteorologiska variabler	Markrelaterade variabler
Lufttemperatur	Yttemperatur
Luftfuktighet	Luftfuktighet, Daggpunktsdeficit
	Frostbeläggningar
	Vattenutfällning i form av dagg
Nederbörds mängd	Nederbörds mängd
	Nederbörd i form av regn
	Nederbörd i form av snö
	Akkumulerad nederbörd i form av regnvatten
	Nysnö (bidrar till snö i terräng)
	Smältande nysnö (bidrar till vatten på marken)
	Akkumulerat snötäckte i terräng
	Smältande snö från terräng
	Avdunstning av vatten från mark
	Vatten på marken
	Is
	Modd
Vindriktning	Vindriktning
Vindstyrka	Vindstyrka
Vindbyar	Vindbyar
	Drivbildning



Figur 5.4: Meteorologiska variabler och markrelaterade variabler.

ackumulerats under de senaste tre timmarna, smältande snö från tidigare ackumulerat snötäckte, mängden nederbörd som avdunstat, vatten på marken som resultat av alla dessa processer.

Samtliga ovan nämnda nya variabler beskriver förlopp i marknivå. Dessa variabler fanns i de prognoser som användes under projektet. I tabellen i Figur 5.4 visas en översikt över två olika grupper av väderrelaterade variabler: meteorologiska variabler, respektive variabler som beskriver mark.

De flesta studier byggde på att ta fram eller testa nya "vädertyper", som representerar unika kategorier definierade med hjälp av en eller flera variabler. Nya kategorier anpassade för och specifika för järnväg kunde föreslås utifrån kunskap som samlas in, härledas utifrån närliggande domäner som vägväder, eller utgöra resultat av kvalitativ analys av data. Vid kvalitativ analys användes systemet ShareWeather som levererar anpassade väderprognoser med en rad variabler enligt 5.4, samt analysverktyget som visualiserar olika diagram, så kallade meteogram, på en tidslinje (5.3). I projektet gjordes anpassningar för att applikationen ska kunna ta emot händelsedata, i syfte att jämföra väder och händelser för samma tidpunkt eller tidsintervall och analog geografisk position eller område.

5.5 Resultat

5.5.1 Dataunderlaget

Vi vet alla att, även om man möjligen skulle önska, vädret inte går att kontrollera. Det är viktigt att poängtera att projektets studier designades, och resultaten sedan tolkades, utifrån tillgång till data och att en del av dataunderlaget är just väder. Ju mer data desto mer tillförlitligt resultat. För ovanliga vädertyper finns mindre data att tillgå och studera, medan de vanligare vädertyperna bjuder på bättre dataunderlag.

Detta innebär att man får flera mätpunkter för en viss vädertyp och/eller orsakskod om man studerar hela landet samtidigt. Men ju mer data man inkluderar för exempelvis ett större geografiskt område (alla Sveriges driftledningsområden), eller samtliga orsakskoder, desto flera beroende variabler finns i dataunderlaget, som position och orsakssamband knutna till specifika orsakskoder som bidrar till att utlösa en händelse.

Men inte bara väder är beroende av bra data. Erfarenheten och den kunskap som samlades in under projektet pekade tydligt på att stora skillnader kan förekomma avseende de lokala förutsättningarna, klimat, geografi och topografi, infrastruktur som exempelvis enkelspår och banans egenskaper och underhållsstatus, trafiktäthet, och inte minst sättet att arbeta inom olika driftledningsområden. Därtill tillkommer andra okända samband och slutligen slumpen.

Resultaten nedan analyseras och diskuteras utifrån dessa utmaningar och perspektiv. Den allmänna slutsatsen är att man i framtida forskning bör utöka dataunderlaget till flera säsonger och år. Men även om man inkluderar ett större underlag återstår vissa utmaningar, exempelvis förekomst av vissa ovanliga vädertyper samt uppdelning i allt större antal vädertyper för att identifiera de kombinationer av variabler som utgör nyckeln till förståelsen av, och framtida prediktioner av, händelser i järnvägssystemet med avseende på väder.

5.5.2 Vädertyper

Inledningsvis användes kategorier kända utifrån dokumenterad kunskap. Dessa var: låga temperaturer (kyla), höga vindstyrkor, nederbördsmängd, nederbörd i form av snö.

Under en förstudie gjordes en kartläggning av vilka vädertyper som fanns representerade under de tidsperioder då många händelser inträffat eller då merförseningen varit omfattande. Förstudien, som använde standardavvikelse för att kartlägga vilka vädertyper som förekommit oftare än andra, tittade på merförseningar som överskred medelvärdet och en standardavvikelse respektive två standardavvikelser, varvid beräkningar gjordes per delsträcka. Resultatet bekräftade i stort sett relevansen av kategorierna listade ovan, undantaget snö som inte studerades här. Därutöver studerades kombinationen kyla och fukt för vilken samband upptäcktes.

Enligt den kunskap som finns sedan tidigare kan "bra" järnvägsväder anses vara ca +10 °C, molnigt och gärna även lite duggregn. Sämre väder utgörs enligt erfarenheterna av: extremt höga eller låga temperaturer, höga vindhastigheter, mycket nederbörd, särskilt snö, samt lövhalka.

Höga temperaturer studerades inte eftersom dessa inte uppträder under höst och vinter. Lövhalka studerades inte specifikt, men viss kunskap kunde samlas in avseende förekomst och dokumentation av lövhalka.

Utöver huvudkategorierna togs flera underkategorier, kombinationer av flera variabler, fram och kom att bli fler under projektets gång. Några exempel är kombination av temperatur och fukt, respektive drivbildning som är kombination av vind och snö i terräng. Olika kategorier bestod av fastställda intervall av värden för en viss variabel, exempelvis temperatur mellan -7 och -15 °C ("lätt kyla"), -15 till -22 °C ("kyla") och < 22 °C ("sträng kyla").

Övriga kategorier, som inledningsvis gick under samlingsnamnet "bra väder", kom att delas upp och preciseras ytterligare under projektets gång, i samband med att den iterativa processen illustrerad i Figur 5.2 genererade mer kunskap om samband mellan vädertyper och händelser.

Temperatur

Kyla är en variabel för vilken det finns dokumenterad kunskap, man har länge varit medveten om att väldigt låga temperaturer orsakar problem i järnvägssystemet. De viktigaste resultaten i denna studie, som för övrigt bekräftade betydelsen av kyla, är att man inte nödvändigtvis behöver ha sträng kyla för att se vissa samband. Påverkan verkar framträda redan vid temperaturer under -10 °C.

Luftfuktighet

Att fukt i kombination med temperatur kan ge upphov till specifika, och ibland besvärliga fenomen, är känt, inte minst från vägväder. Fukt är en förutsättning för att frost ska fällas ut ur luften och fastna på mark och fysiska föremål. Is, ett annat problem, uppstår när vatten fryser till is. Kunskapsinsamling från operativ personal pekade även på att problem kan transporteras från en punkt till en annan när den omgivande luftens temperatur ändras, exempelvis kan snöbeklädda fordon flytta med sig snö- och ismassor till nya platser med högre temperatur (till lägre terräng eller sydligare breddgrader). När snö och is faller av tåget kan detta orsaka problem med spårväxlar. Fukt är en variabel som kan spela med i alla dessa fenomen och har avgörande betydelse för när problemet uppkommer initialt. Exempelvis påverkar luftfuktigheten hur och om snön lägger sig och fastnar på tåg och exempelvis elkontakter. Luftfuktighet kan beskrivas med exempelvis relativ luftfuktighet, eller m.h.a. daggpunktsdeficit som är differensen mellan daggpunkt och ytemperatur.

Kombination av olika temperaturintervall och fukt testades därför. Samband upptäcktes vid olika studier på orsakskod- och driftledningsområdesnivå och medelvärdet för landet och alla händelser uppvisade exempelvis ett mönster att fukt i kombination med samma temperaturintervall ger flera händelser än torrt inom samma temperaturintervall. Mer om skillnader mellan olika orsakskoder följer i senare avsnitt.

Vind

Vind anses ha stor påverkan på järnvägssystemet. Utmaningen när man studerar påverkan vid höga vindhastigheter är att dessa situationer inträffar relativt sällan. I kombination med snö i terräng kan dock drivbildning påverka även vid lägre hastigheter. Tar man lärdom från andra domäner som vägtransport och flyg kan vindar på 5 m/s orsaka problem.

I detta projekt användes både vindstyrka och vindbyar, med fokus på den senare. Vindriktning borde inkluderas i framtida arbete då man bör dela in vinden i komponenter längs med och tvärs banans riktning. Detta rymdes inte inom ramen för detta projekt, men kan rekommenderas för fortsatt forskning, särskilt som den kartläggning av järnvägsnätet som gjordes lämnar öppet för vidareutvecklingar av denna typ.

Nederbörd

Järnvägen anses, enligt tidigare kunskap, påverkas vid lite större nederbördsmängder, särskilt snöfall, medan små regnmängder till och med anses vara till fördel för järnvägstrafiken. I projektet antogs därför att nederbördsmängder som 0.3 mm till 1.0 mm under tre timmar inte skulle ha någon nämnvärd påverkan. Projektet ville därför fokusera på större mängder, dock var förekomsten av kraftig nederbörd begränsad vilket ledde till en indelning enligt: 0 - 0.3 mm, 0.3 - 1 mm, 1 - 2 mm, 2 - 3 mm, samt > 3 mm.

Tryck

Tryck ingår inte i VViS och fick därför inte utrymme i projektet i stort eftersom de AI-relaterade studier som bedrevs parallellt utgick ifrån VViS och därmed inte inkluderade tryckdata. Variabeln tryck studerades ändå delvis, varvid den visade sig kunna generera en del nyttig information. Lufttryck över, respektive under, normalvärdet som är 1013.25 hPa, testades i flera vädertyper. Tester gjordes även för lägre tryck som exempelvis <1000 hPa och än lägre.

Resultat som är värda att lyfta fram är att vid tryck lägre än normalt, fuktigt och -7 till +3 grader, i samtliga driftledningsområden sammanföll med flera händelser än samma temperaturintervall, högre tryck än normalt med mindre fukt. En naturlig förklaring är att temperaturer kring noll i kombination med lågtryck kan förknippas med flera fenomen som kan orsaka specifika förhållanden som frost och is, och därmed även problem för järnvägen. Därutöver kan även höga vindhastigheter förknippas med lågtryck. Samtidigt fick man dock fram ett nära motsatt resultat för något högre temperaturer inom plusområdet, närmare bestämt > +3, där lägre tryck istället gav färre händelser än högre tryck. Det senare resultatet är av särskilt intresse eftersom det öppnar upp för ytterligare indelningar av vädertyper, som kan användas i framtida forskning för kartläggning av samband mellan väder och händelser i järnvägssystemet.

Vi kan därför dra slutsatsen att variabeln tryck kan vara ett bra komplement och att nya samband kan upptäckas genom att introducera en kombination av tryck och andra variabler. Vid analyser inom synoptisk meteorologi används ofta tryck för att definiera vanligt förekommande vädermönster och tillämpa kategoriseringar liknande dem som gjorts i detta projekt genom definition av "vädertyper". Man bör härmed tillägga att det för "järnvägsväder"

behövs betydligt flera "vädertyper" än för ett flertal andra meteorologiska applikationer som exempelvis vindkraft eller spridning av luftföroreningar, vilket i sig understryker att järnvägssystemet verkligen är komplext.

Bra väder

Lika viktigt som att kartlägga och studera de vädertyper som ger störst påverkan, är att definiera vad som är "bäst" ur järnvägssystemsperspektiv.

Uppdelning i intervaller med avseende på både temperatur och andra variabler som fukt och tryck, kom att generera flera nya intressanta resultat. Slutligen fastställdes utifrån resultaten att "bra väder" är avgränsat till -7 till +3 grader, torrt och stabilt, med endast mycket små tillåtna mängder nederbörd, respektive nysnö och snö i terräng (snödjup i terräng sattes till < 0.5 cm för att inte minska datamängden för mycket), samt vindbyar < 15 m/s.

När en liknande vädertyp definierades, innehållande samma villkor men ett mildare temperaturintervall, > +3, fick man resultat som avvek, vilket tyder på flera fenomen som uppträder under "bra" väder, varför framtida forskning bör arbeta med att identifiera flera vädertyper som inte faller inom extrema intervall.

Lövhalka

Antalet fall med spårhalka (IFK01) var begränsat i relation till andra väderfenomen och innehöll, utöver lövhalka, situationer med halka orsakade av vinterförhållanden. Under perioden 23/10 - 15/11 registrerades endast två trafikpåverkande händelser klassificerade som IFK01 (spårhalka) och gällande lövhalka. Analys av text gav dock att lövhalka kan förekomma under andra orsakskoder än IFK01. Några exempel är: naturhändelser ONA- där händelse H3315069 fått beskrivningen "Spårhalka mellan Habo-Mullsjö" och händelse H3322418 innehöll texten "Spårhalka spår 5...", samt signalfel ISA03 händelse H3310402 innehållande texten "halt på spåren". Samtliga händelser inträffade under perioden 23/10 - 5/11 2020.

Fortsatt arbete borde utforska tidiga höstmånader, varvid kvalitativ analys borde ingå för att upptäcka händelser kopplade till lövhalka. En annan rekommendation är att ta fram en vädertyp som tar hänsyn till hur temperaturer och fukt sett ut under föregående vecka, för att utöka förståelsen för lövhalka. I de prognoser som använts fanns även variabeln växthastighet som beskriver vegetationen.

5.5.3 Samband mellan väder och händelser

Generellt om variabler associerade med händelser

Projektet kom att fokusera på två variabler, trafikpåverkans längd respektive händelsefrekvens, varav båda uppvisade samband med vissa vädertyper. Det kanske viktigaste resultatet av studien av samband mellan väder och händelser i järnvägsnätet är just att flera händelser verkar inträffa vid vissa vädertyper, men att förseningar och trafikpåverkans längd troligen påverkas av en del andra faktorer som vädret inte direkt rör över.

Resultaten pekade på att det troligen finns samband mellan å ena sidan nederbörd, vind, kyla, respektive kyla och fukt, och å andra sidan merförsening. För händelsens varaktighet var sambanden inte tydliga. Detta kan förklaras med att händelseregistrering inte sker fullt strukturerat. En rad liknande studier gjordes med andra vädertyper och för flera månader och även med en uppdelning av landet i söder och norr, innan fokus kom att läggas på trafikpåverkans längd och händelsefrekvens. Studierna av händelsefrekvens fördjupades efterhand mot driftledningsområden och orsakskoder.

Ett resultat är nämnvärt i detta sammanhang och det avser mänsklig perception och hur händelsefrekvens står i relation till de andra variablerna som beskriver störningens natur. Om många händelser inträffar under exempelvis 10 timmar, kommer olika intressenter alltifrån operativ personal och samarbetspartners till slutkunder och resenärer att uppleva den sammanlagda påverkan och inte enbart enstaka förseningar. Summan av exempelvis alla merförseningar eller summan av all trafikpåverkan i minuter kan bli stor om betydligt flera, låt oss säga dubbelt så många, händelser inträffar, även om merförseningen eller trafikpåverkan per händelse inte varit större än vanligt. Hur man, i rollen som resenär eller operativ personal, upplever störningar kan vara relaterat till perceptionen av summan av alla förekommande störningar man fått information om under flera timmar.

Trafikpåverkan

Resultaten visar att det är svårt att påvisa samband mellan vädertyperna och trafikpåverkans längd, även om mönster uppträder för flera vädertyper. Detta blir än mer tydligt när man går ner på orsakskodnivå 3. Särkilt uppmärksamhet bör också ägnas åt geografisk position, eftersom förutsättningarna kan se väldigt olika ut för olika järnvägssträckor. Vidare kan trafikpåverkans längd påverkas märkbart om man, på grund av väderinformation man erhållit, börjar agera för att reducera påverkan. Trafikpåverkan som variabel och som den är utformad idag, verkar vara för komplex för att man ska kunna fastställa relationer mellan vädertyper och trafikpåverkans längd. En alternativ tolkning är att vädret inte påverkar just denna variabel.

Dock finns några undantag och anledning att diskutera resultaten. Undantagen finns bland extremvärden. I Figur 5.5 visualiseras trafikpåverkans längd m.h.a. cirklar runt händelsens position. Cirklarnas färg motsvarar en viss grupp av vädertyper (blått står för nederbörd, gult och orange för vind, ljusgrönt och vitt för frost och is, violett för kyla, medan bra väder är representerat i grönt). Extrema värden av vind (kraftig blåst) och temperatur (sträng kyla), som redan i inledningen identifierats som intressanta vädertyper, inträffar relativt sällan. Medan resultaten i Figur 5.5 delvis bekräftar samband mellan trafikpåverkans längd och kyla respektive kraftiga vindar, fanns inga tydliga samband för mer normala vinterförhållanden som exempelvis temperaturer mellan -7 och -15 grader.

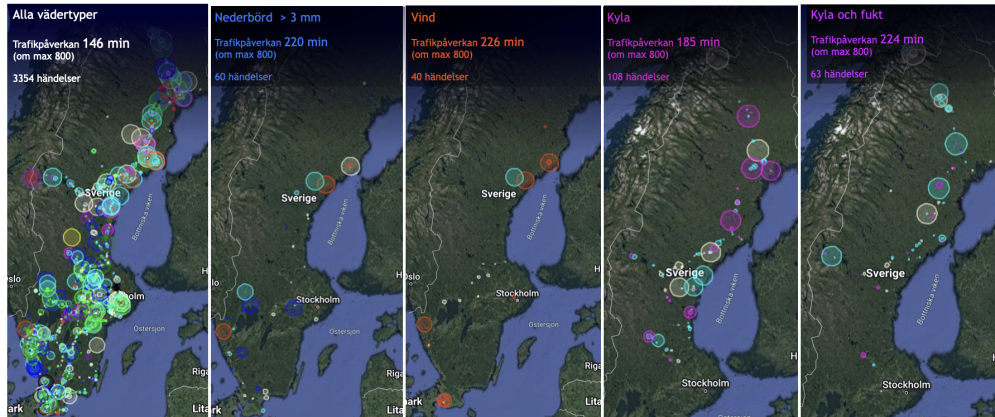
Att mycket låga temperaturer visade sig sammanfalla med större trafikpåverkan än under mer normala förhållanden syns även i SHAP-analysen som presenterades i Figur 3.16, där låga ytemperaturer visserligen verkar påverka prediktionen av trafikpåverkan, men på grund av att kyla (låga ytemperaturer) inträffar sällan (se Figur 3.11), är mönster svåra att upptäcka m.h.a. AI-baserade metoder. AI-modelleringen pekade, enligt tidigare avsnitt, på att nederbördstyp och ytemperatur påverkar trafikpåverkanslängden något men inte i någon större omfattning. Under fokus väder kunde samband, som inte framträder i AI-modellen, upptäckas. Resultaten är dock osäkra även när kvalitativ analys användes som komplement för att tränga djupare in i problemet. Eftersom extrema värden är ovanliga och resultaten inte entydiga, måste slutsatserna här bli generellt vaga.

En möjlig förklaring, och i analogi med resonemanget bakom Figur 3.15, är att andra faktorer påverkar trafikpåverkans längd mer än väder. Här följer ett exempel som illustrerar detta. Utöver att de är ovanliga, är händelser med sträng kyla (< -22) koncentrerade till ett driftledningsområde (Malmbanan i Boden) i norra Sverige. De norra driftledningsområdena har olika förutsättningar och infrastruktur med exempelvis enkelspår och mer svårtillgänglig terräng. Enligt Figur 3.9 är Boden det driftledningsområde som har längst trafikpåverkan i relation till de övriga driftområdena. Även för intervallet -15 till -22 grader finns relativt många händelser i norra Sverige, medan Götaland och södra Svealand är svagt representerade. För intervallet -7 till -15 (och fukt) finns en del dataunderlag även för södra Sverige. Trafikpåverkans längd visade sig generellt vara kortare i södra Sverige (129 min) än i norra Sverige (198 min) samtidigt som händelser inträffar oftare i söder (2561 händelser) än i norr (816). För intervallet -7 till -15 ligger trafikpåverkan nära medel (128 min i södra Sverige respektive 186 min i norr). Vi kan nu dra slutsatsen att, eftersom dataunderlaget för sträng kyla är baserat på enbart norra Sverige, kan resultatet bero på skillnader mellan norra och södra Sverige, snarare än väder. Därför måste vi studera driftledningsområden var för sig eller på annat sätt särskilja olika delsträckor. En bra början är att dela in Sverige i en nord och syd. Medan resultaten här presenteras som medelvärden för norra respektive södra delen av landet, utgår analysen ifrån resultaten för samtliga vädertyper per driftledningsområde.

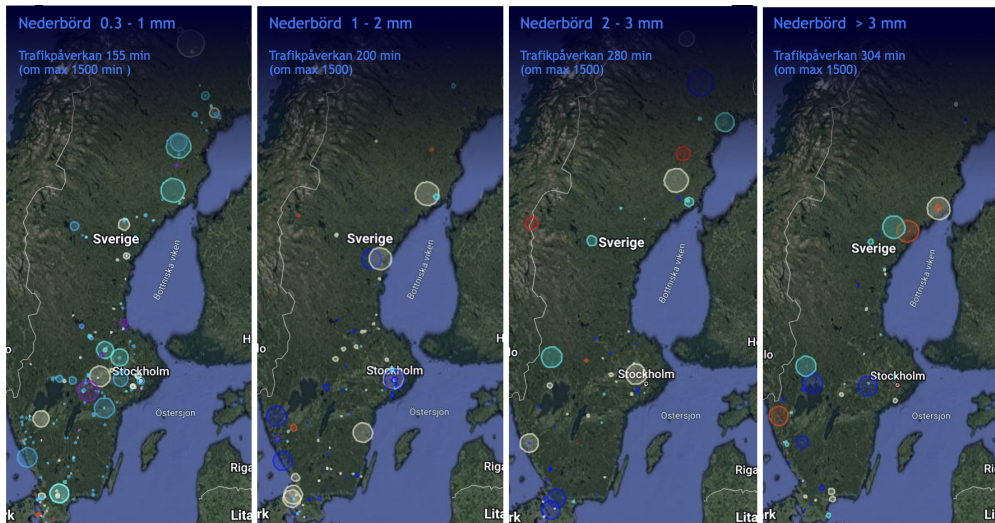
Resultaten för vind (se 5.5) och nederbörd (se 5.6) är mer tillförlitliga. Situationer med vindbyar över 18 m/s var lika vanliga i söder som i norr. Resultaten visar att trafikpåverkan ökade med ökad vind. Liknande resultat erhöles för ökad nederbördsmängd samt kombination av kyla och fukt.

Den generella utmaningen med variabeln trafikpåverkans längd var dess fördelning. Det aritmetiska medelvärdet var 146 minuter för den studerade perioden (20201023 - 20210228), efter att en övre gräns på 800 minuter införts. Alla värden över 800 minuter tilldelades värdet 800. Värdena var asymmetriskt fördelade med stora positiva avvikelser. Medelavvikelseerna var -50 minuter respektive +416 minuter. Om man betraktar det aritmetiska medelvärdet för respektive vädertyp kan dock mönster upptäckas, som i Figur 5.5 och 5.6.

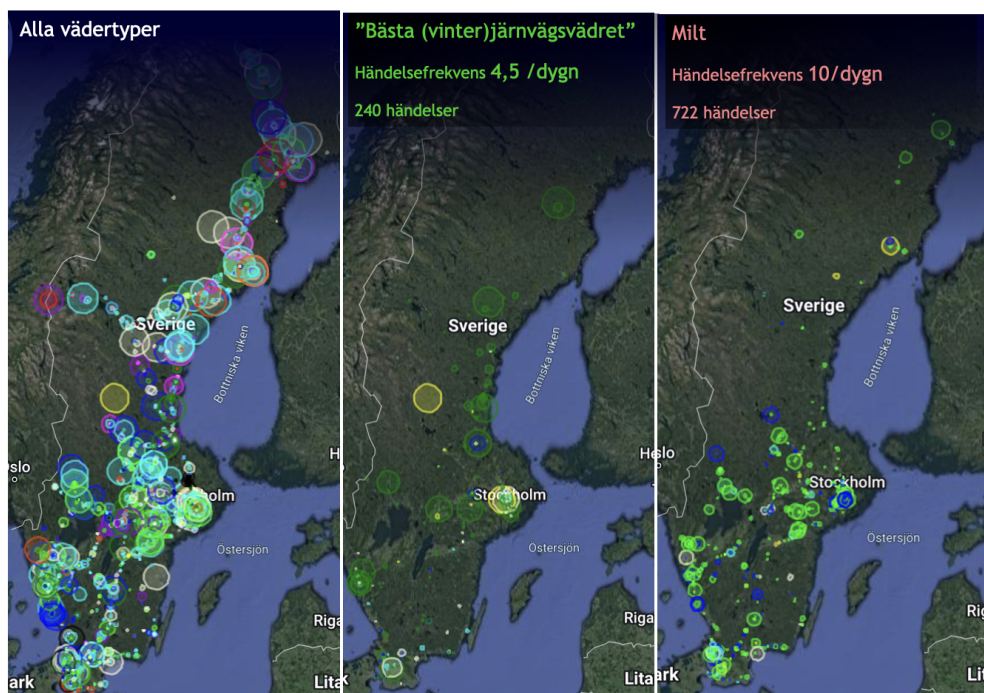
De lokala skillnaderna upphör dock inte vid gränsen mellan driftledningsområdena. Inom driftledningsområdena i södra Sverige finns olika typer av järnvägsinfrastruktur och framför allt olika trafiktäthet med koncentration till storstadsregionerna. I Figur 5.7, som visualiserar trafikpåverkans längd för några vädertyper som kan betecknas som "bra" väder, kan man se att många händelser med stor trafikpåverkan vid "bäst väder" - en vädertyp som identifierades utifrån lägsta händelsefrekvens - inträffat i närheten inom storstadsregionerna, samtidigt som många andra händelser i inlandet har kort trafikpåverkan. Resultatet för landets södra respektive norra del är en något högre medeltrafikpåverkan (140 respektive 241 minuter för söder respektive norr) än genomsnittet (129 respektive



Figur 5.5: Trafikpåverkans längd visualiserad som radiens storlek runt positionen för händelser.



Figur 5.6: Trafikpåverkans längd vid olika nederbördsmängd.



Figur 5.7: Vädertyperna med bra väder, Bäst respektive Milt.

198 minuter), men baserat på visualiseringen i kartan i Figur 5.7, kan man dra slutsatsen att det kan finnas stora skillnader inom ett driftledningsområde i sig. Den tidigare analysen av de data som användes i AI-modelleringen i delprojektet Prediktion av tidsåtgång för trafikstörande händelser, baserad på ett större dataset (se Figur 3.9), visar att Stockholms driftledningsområde har en lägre trafikpåverkans längd per händelse i jämförelse med de övriga driftledningsområdena. Resultaten från studierna med fokus på väder och samband mellan väder och händelser i järnvägsnätet var i stora drag analogi med detta resultat, dock upptäcktes intressanta skillnader. Även här (se Figur 5.7) hade Stockholms och därutöver även Göteborgs driftledningsområde kortast trafikpåverkans längd totalt sett, men inom ett och samma driftledningsområde framträder några geografiska skillnader som kan vara systematiska. Inom en smal radie runt storstadsområdena inträffade händelser med generellt större trafikpåverkan. Detta är tydligast för vädertypen "bäst väder". Närmare bestämt hade de (45) händelser som inträffat inom en radie av 50 km runt Stockholm en trafikpåverkan på 176 minuter och Göteborg 180 minuter och 28 händelser.

En fördjupning i skillnader inom driftområden och storstadsregioner ger att storstadsregionerna, i kontrast till det förra resultatet för "bäst väder", oftast har kortare trafikpåverkan än hela driftområdet för vädertyper som har påverkan på både trafikpåverkans längd och händelsefrekvens. Vid vädertypen fukt och -7 till -15 °C inträffade proportionerligt många (103) händelser inom 50 km radie från Stockholm med i genomsnitt 103 minuters trafikpåverkan. Detta resultat kan tillskrivas slumpen, men det kan finnas andra förklaringar som hänger samman med storstadsregionens unika situation relaterad till exempelvis pendeltågstrafiken som möjligen påverkas av andra tröskelvärden än den övriga spårtrafiken.

Slutligen bör betydelsen av händelsens orsak diskuteras. Analys av trafikpåverkans längd per vädertyp och orsakskod, där interntexter utgjorde visst stöd, pekade på att trafikpåverkans längd kan variera avsevärt beroende av orsakskod. (Detta resultat stöds även av den analys av trafikpåverkan per orsakskod som gjordes under AI-studierna 3.7(b) baserad på ett betydligt, ca 10 gånger, större dataunderlag). Efter den förstudie som i början av projektet gjordes med fokus på orsakskoder med direkt koppling till väder, som IFK med spårhalka, snö och is (trafikpåverkan 261 minuter), samt naturhändelser ONA (320 minuter), kunde man se att händelser associerade med dåligt väder hade en märkbart större trafikpåverkan. Men även andra orsakskoder kan ha direkt koppling till väder utan att det framgår av kodningen. Analys av manuellt skapade texter som beskriver händelsen, bland annat med specifika väderrelaterade sökord som "blåst", "oväder", "halt" gav visst stöd för att händelser som inträffar vid lite "sämre" väder tenderar att ha längre trafikpåverkan.

Det kan även finnas specifika orsakskoder som inträffar mer sällan, är känsliga för specifika vädertyper, och

samtidigt har längre trafikpåverkan. Ett exempel är IEA01 Kontaktledning, som har längre trafikpåverkan än de flesta övriga orsakskoder. Trafikpåverkans längd för IEA01 var 328 minuter i norra Sverige och 308 minuter i södra delen av landet. Andra orsakskoder på orsakskodnivå 3 som utmärker sig är ISA01 (medan gruppen ISA, d.v.s. orsakskodnivå 2, totalt sett har lägre trafikpåverkan) och IBÖ01.

Sammanfattningsvis kunde samband mellan väder och trafikpåverkans längd upptäckas för inte fullt så extrema värden. Exempelvis verkar nederbörd på 2 eller 3 mm under de senaste 3 timmarna ge en något längre trafikpåverkan. Likaså ger temperaturer under -15 °C en tydlig ökning av trafikpåverkan, medan det möjligen kan finnas samband för temperaturer mellan -7 och -15 °C särskilt om det samtidigt är fuktigt. Fukt i kombination med kyla gav större påverkan än endast kyla. Vind ger en tydlig ökning vid vindbyar över 18 m/s , men för vissa orsakskoder kunde samband ses för lägre intervall. För vindbyar $> 10\text{ m/s}$ i kombination med snö i terräng var påverkan relativt stor i norra Sverige. Generellt är alla dessa samband svåra att studera med anledning av trafikpåverkans egenskaper som variabel, bl.a. den stora spridningen. Utifrån resultaten presenterade i Figur 3.16, där några vädervariablers påverkan på AI-modellen presenteras, finns samband mellan trafikpåverkans längd och: yttemperatur (både låga värden, d.v.s. kyla, och framför allt höga värden, som inte studerades här), daggpunktsdeficit (låga värden, vilket betyder hög fuktighet och risk för exempelvis frostbeläggningar), nederbördsmängd. Detta ligger i linje med de resultat som framkom vid studier av samband mellan väder och händelser i järnvägsnätet för perioden 20201023 - 20210228 under fokus väder. Några variabler avser summering av värden över flera dagar, exempelvis 3-dygnsmedelvärden för yttemperaturen respektive nederbörd, varav det framgår att låga temperaturer under några dygn eller nederbörd under några dygn medför något större trafikpåverkan. Kvalitativ analys av väder och händelser på en tidslinje visade också på samband av denna typ. Flera fenomen som förknippas med exempelvis besvärligt väglag är resultat av processer som kan ta timmar eller dygn, exempelvis is. Det är naturligt att anta att även järnvägssystemet ställs inför liknande utmaningar, varför vidareutveckling av nya vädertyper som bygger på flera variabler och beräkningar av konsekvenserna på marken, snarare än meteorologiska variabler, rekommenderas.

Händelsefrekvens

Projektets viktigaste resultat från fokusområde väder är samband mellan olika vädertyper och händelsefrekvensen. Vid vissa vädertyper verkar flera händelser i järnvägssystemet inträffa och sambandet kunde kvantifieras för den tidsperiod som studerades. Resultaten är sammanfattade i Figur 5.8.

De vädertyper som enligt resultaten har störst påverkan bekräftar tidigare erfarenheter och antaganden, exempelvis kända samband avseende kyla och vind. Dock tillkommer en del kunskap avseende nya kombinationer av variabler. Ett exempel är drivbildning som är kombination av vindstyrka och ackumulerad snö i terräng, ett annat kombination av kyla och fukt.

Dock är det inte enbart extrema värden som studerats under projektet och genererat nya resultat. Sedan tidigare har man haft kunskap om vad "dåligt" väder innebär och vilka vädervariabler som brukar vara inblandade. Men utöver "dåligt" väder finns en skala av vädertyper mellan "dåligt" och det "bästa" vädret. Här har projektet haft ambitionen att ge en helhetsbild samtidigt som det utforskat nya mer avancerade kombinationer av variabler och upptäckt nya samband.

Eftersom geografiska skillnader finns avseende naturlig förekomst av vädertyper, d.v.s. klimat, och de regionala skillnaderna dessutom är omfattande (trafiktäthet, infrastruktur som exempelvis enkelspår, samt underhåll vilket speglas av baskontrakt) kan resultatet studeras områdesvis. Figur 5.8 sammanfattar resultaten för norra (Boden, Ånge, Gävle) respektive södra Sveriges (Hallsberg, Stockholm, Norrköping, Göteborg, Malmö) driftledningsområden.

Enskilda områdets avvikelse från landsmedelvärdet och medelabsolutavvikelsen mellan olika driftledningsområden visar att resultaten är godtagbara även när man slår ihop områden, dock med några undantag. Den procentuella medelabsolutavvikelsen (MAPD), varierade i de flesta fall mellan 10 och 20 procent. För de lägsta temperaturintervallen ($< 22\text{ °C}$) var den 25 procent, medan kraftiga vindbyar (56 procent), drivbildning (38 procent), marktyp med nysnö hade avvikande värden. Att kraftig vind hade stor spridning kan förklaras med att detta väderfenomen inträffar mer sällan och framför allt kan slå mycket lokalt, samtidigt som de har stor påverkan. Nederbörd $> 3\text{ mm}$ hade MAPD 43 procent vilket kan förklaras med att alla områden inte fick lika många tillfällen med större nederbördsmängd, samtidigt som lägre nederbördsmängder $< 3\text{ mm}$ var betydligt vanligare och hade MAPD mellan 12 och 20 procent. Kyla med temperaturer inom intervallet -7 till -15 °C hade också en stor spridning trots en bra datarepresentation, vilket antyder att det inte finns ett starkt samband mellan just detta intervall och

händelsefrekvens, även om intressanta resultat kunde fås för enskilda orsakskoder.

Kyla

Vid sidan om att resultaten bekräftade betydelsen av stäng kyla med låga temperaturer som löper ner mot -20 °C och därunder, har projektet studerat mildare temperaturer. Ett resultat är att vissa typer av problem kan uppstå redan vid temperaturer runt -10. För temperaturer under -15 °C gav projektets studier att det finns en påverkan.

Luftfuktighet

Ny kunskap som projektet bidrar med gäller kombination av låga temperaturer och fukt, där fukt och samma temperaturintervall gav flera händelser än samma temperaturintervall och lägre fuktighet. Exempelvis gav medelvärdet för hela Sverige 17 respektive 11 händelser per dygn för samma temperaturintervall -7 till -15 °C. Marktyp frost, som var tämligen välrepresenterad som vädertyp, gav möjligen inte det förväntade resultatet, utan hade en mer begränsad inverkan. Dock kan resultatet bero på den definition av frost som marktyp och den hierarkiska ordningen som infördes marktyperna emellan, då flera marktyper kan råda samtidigt (exempelvis kan det snöa eller finnas isbeläggning på marken samtidigt som det är frost som ju bildas när det är fuktigt i luften och samtidigt tillräckligt kallt).

Vind

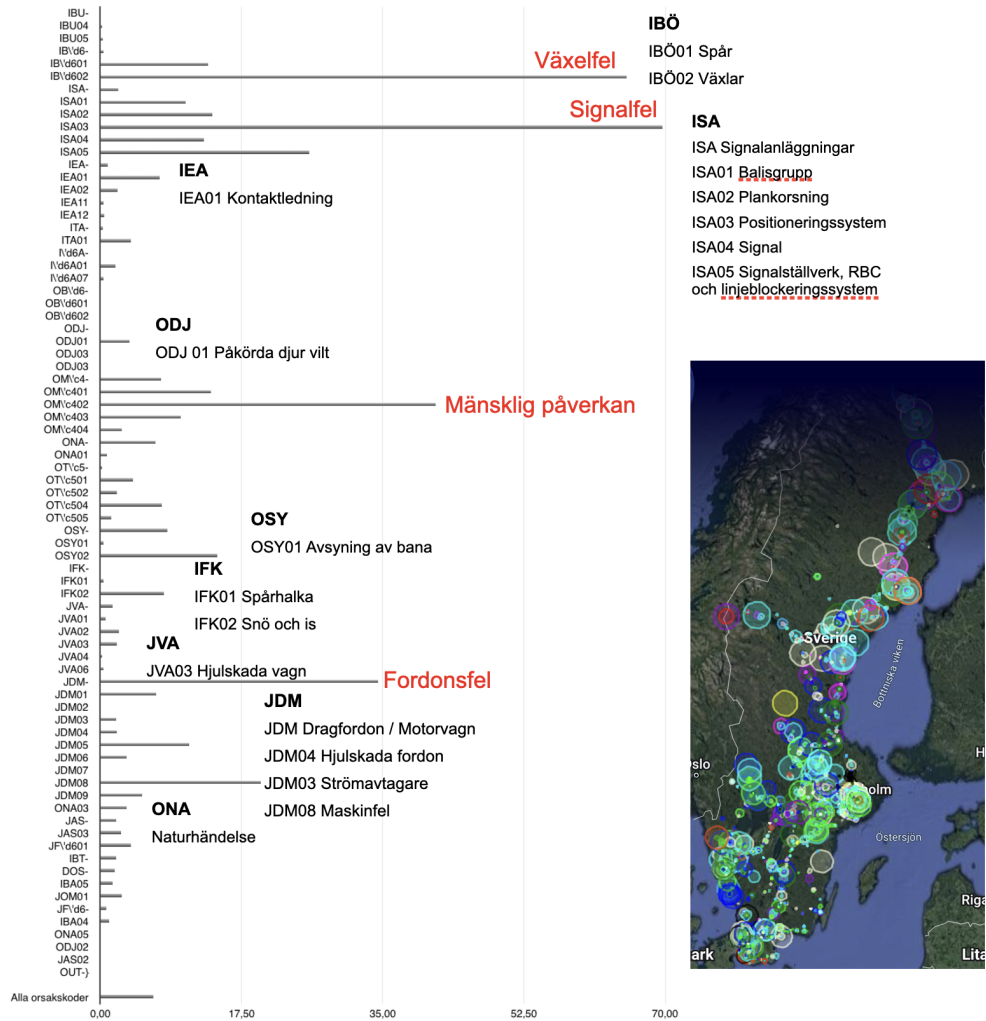
Blåsigt väder är ett känt problem vilket resultaten bekräftade. Vindbyar över 18 m/s (vilket ungefär motsvarar en medelvindhastighet på över 10 m/s) gav mycket tydlig ökning av händelsefrekvensen, men i vissa resultat - områdesvis och per orsakskod - kunde påverkan ses redan vid 15-18 m/s. Drivbildning, som definierades som vindbyar över 10 m/s samtidigt som snö i terräng överskred 0.5 cm snö, gav också en påfallande ökning av händelsefrekvensen.

Nederbörd

I fallet nederbörd noterades överlag en ökning av antalet händelser vid ökad nederbördsmängd. Detta mönster var även, med ett par undantag (Boden och Malmö) tydligt för olika driftledningsområden samt när olika månader studerades var för sig.

Geografiska skillnader

Vissa orsakskoder inträffade oftare i norra Sverige i relation till andra orsakskoder, vilket framgår av Figur 5.9. När man studerar antal händelser per driftledningsområde och orsakskod kan man även se att följande orsakskoder varit ovanligt frekventa i norra Sverige: IBÖ02 (Boden, Ånge), JDM08 (Ånge, Gävle, Hallsberg), OTÅ04, ONA-, ODJ01. I södra Sverige inträffade följande orsakskoder oftare i relation till andra orsakskoder: JDM-, OMÄ02, IBÖ01, ISA01 (Malmö), ISA02 (Malmö). Signalfel av typen ISA01 och ISA02 liksom människor i spår OMÄ02 kan förklaras med mer komplex infrastruktur med tätare bebyggelse (flera plankorsningar och större befolkningensmängd). Dessutom pekar resultaten på att ISA01 och ISA02 inte är beroende av låga temperaturer som ju företrädesvis råder i norr. Likaså är det rimligt att norra Sverige är hårdare drabbat av viltolyckor och kallare väder som ger problem med växlar (IBÖ02 visade ett tydligt beroende av kyla).

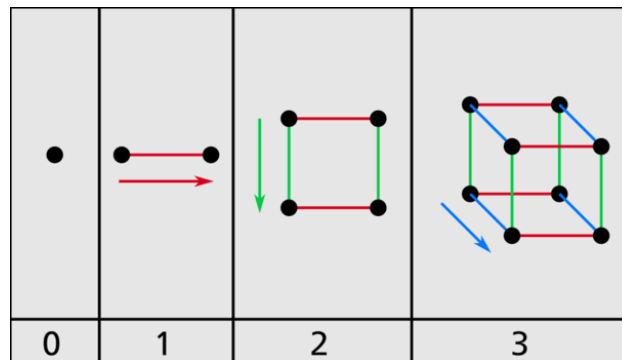


Figur 5.9: Förekomst av orsakskoder 20201023 - 20210228.

Kapitel 6

Visualisering av trafiken och händelser i 3D

Visualisering är en viktig del för att överföra information till användaren som kan tolkas och analyseras för ökad förståelse. Vid användning av 3D (tredimensionell visualiserings teknik) ökar möjligheten till informations överföring om man på rätt sätt strukturerar informationen.



Figur 6.1: Visualiserings koncept 0D till 3D visualisering

Med datoriseringens utveckling har det varit möjligt underlång tid att utforma visualisering i 3D, i början mer enskilda objekt och med spelsektorns utveckling har behovet av kraftigare datorer, både CPU (Central Processor Unit) och GPU (Graphical Processor Unit) har möjliggjort att mer kraftfulla och mer detaljerade visualiseringar är möjliga.

The Matrix Awakens är en ny generation av digital verklighetsrealism som inte setts tidigare men de nya funktionerna som nu erövrar spelvärlden bygger 3D system realiserade i Unreal Engine 5.

Unreal Engine 5 är en digital grafisk motor för alla olika typer av dator baserad 3D visualisering som genom sin release April 2022 med möjligheten till stora världskordinater, "eng. Large World Coordinates (LWC)"

Detta gör att det är möjligt att spänna upp hela Norden i en hög upplöst modell. Järnvägsnätet har tagits från öppen data som man kan se gäller främst Sverige men även vissa sträckor i de övriga nordiska länderna.

För lanskapet användes en modell som heter Cesium som med markerade trafikplatser och järnvägs sträckor i ett Järnvägsnät i ett Unreal Engine 5. Utvecklingen av systemet är framtaget av AIRT i samarbete med AFRY.

Inom AIRT definierar vi det som ett rapid development dvs. ett snabbt utvecklings uppdrag där AFRY utvecklat systemet i Unreal Engine 5

Målet med arbetet är att påvisa möjligheter med visualiseringar i 3D miljöer. Prototypen som är utvecklad är inte riktad till någon specifik roll utan syftar till att påvisa framtida möjligheter kring visualisering.

Det handlar om att visa analys resultatet från AIRT som tex. tågensrörelser med ankomst och avgång från stationer och information om trafikstörande händelser prognoser kopplade till järnvägssystemet.

Sambandet mellan trafikstörande händelser och tågens förmåga att komma i tid har stark system beroende. Trafikstörande händelser påverkar tåg som påverkar andra tåg, i fler led.

Detta med komplexa samband som är svåra att identifiera men kan visualiseras med hjälp av 3D tekniken.



Figur 6.2: Unreal Engine 5 The Matrix Awakens



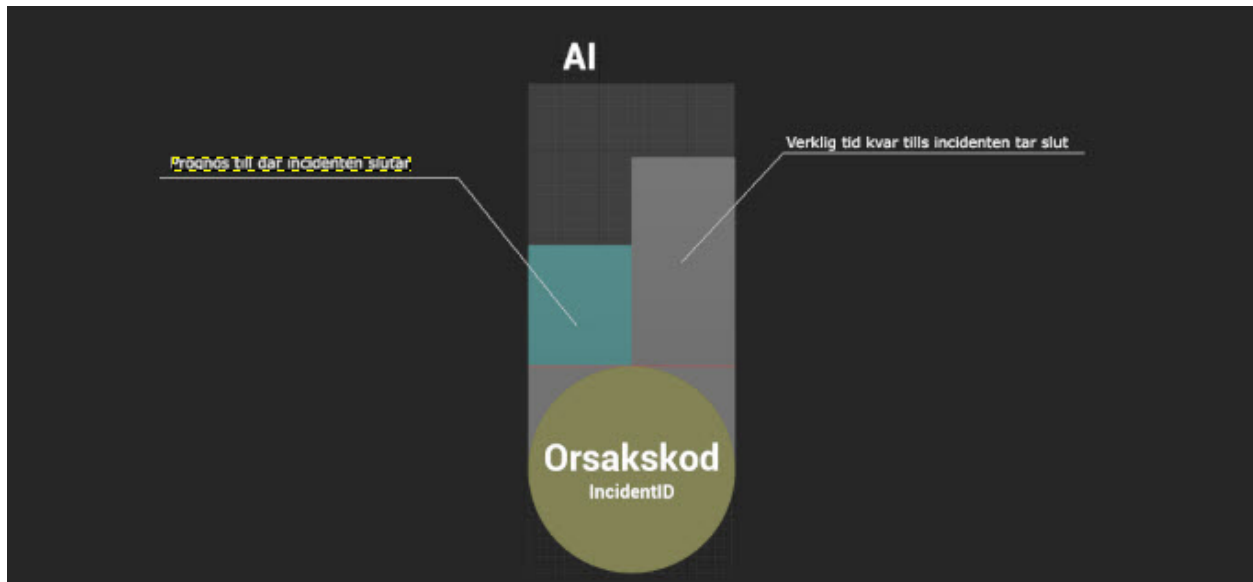
Figur 6.3: Unreal Engine baserad visualisering

Resultaten för trafikstörande händelser visas som prognos och visas som visualisering på platsen där de trafikstörande händelsen rapporterats.

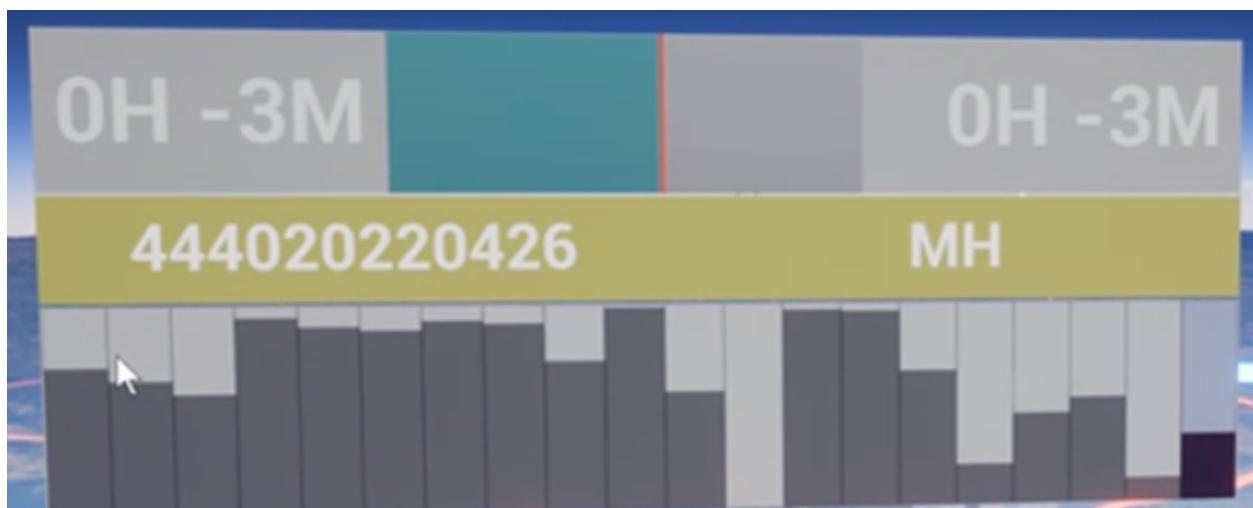
Där det står Orsakskod och incident ID som är händelse id i data som visas i den runda cirkeln nederst. På den högradelen "Verklig tid kvar tills incident tar slut" indikerat höjden på stapeln i grå färg. På den vänstradelen visas prognos från AI modelens resultat.

Tåg i tid ser annorlunda ut där botten fältet visar en stapel graf över alla stationer som tåget kommer besöka och hur nära verkligheten prognosen ligger. En låg stapel visar bättre prognos. Mörkaste färg visar aktuell station. Vid tom stapel är det indikation på saknad data.

Baren i mitten innehåller UppdragsID till vänster som är sammansatt av tåg id och datum i exemplet 4440 som tåg id och datum 20220426. Till höger identiteten för nästa Trafik plats.



Figur 6.4: Trafikstörande Händelser prognos visualisering



Figur 6.5: Tåg i tid prognos visualisering

Överst till höger visas antal minuter tid kvar till nästa station enligt prognosen. Överst till vänster tid kvar till station enligt verklig ankomst tid.

Överst i mitten kring det röda strecket indikeras prognos och verklig tid.

Översikt som del av Sverige kring Göteborg som en del av västra Sveriges järnvägs sträckor. Tåg i tid data är presenterat på sträckan Göteborg till Stockholm , här i bilden fram till ser vi Skövde och lite till.

I bilden har vi två händelser mellan Göteborg och Säffle. Vi ser totalt sex tåg på sträckan Göteborg till norr om Skövde.



Figur 6.6: Del av Sverige som visar prognos visualisering

Kapitel 7

Diskussion

7.1 Förslag på hur prediktioner kan integreras i verktyg

Då en händelse väntas ha en trafikpåverkan behöver den uppskattas i verktyget Operativt Lednings System, OLS. Användaren har för nuvarande stöd i form av en uppskattning baserad på orsakskod samt driftledningsområde och visas i form av tid illustrerad i figur 7.1a. För att stödja användaren föreslår vi att mer information integreras i verktyget. Denna information syftar till att dels ge operatören mer information om hur säker modellens prediktion är, och dels mer specifik information om varför modellen har gjort just den slutsatsen. För att stödja i förståelse av modellens säkerhet kan konfidensintervall visas, se figur 7.1b, eller fördelningen av liknande händelsers trafikpåverkan, se figur 7.1c. För att få mer specifik information kan vi visa hur mycket olika variabler har påverkat modellens prediktion, se figur 7.1d.

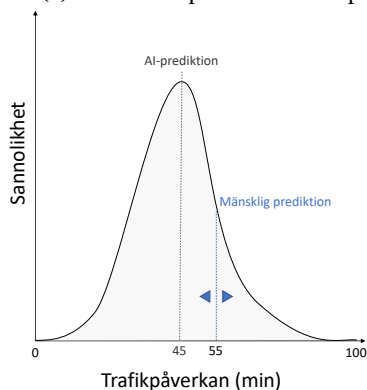
I figurerna 7.1c och 7.1d kan interaktion integreras genom att operatören drar i reglagen. Sådan interaktion gör att operatören enklare kan förhålla sin kunskap om situationen i relation till modellen.

För att se till att operatören inte tappar förmåga i att uppskatta trafikpåverkan så kan den anges i två steg. Först utan AI-stöd och sedan med AI-stöd. Detta förfarande kan även ligga till grund för att utvärdera hur mycket beslutstödet hjälper operatören.

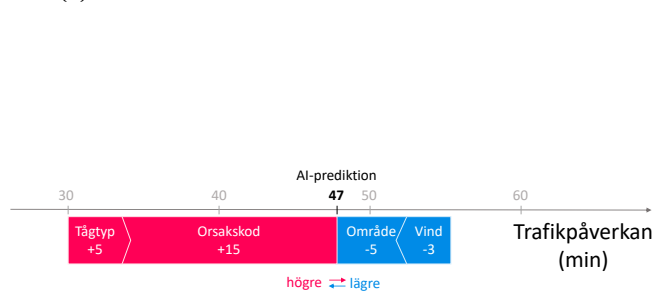
47 min

47±5 min

(a) Nuvarande presentation av prediktion.



(b) Prediktion inkl. konfidensintervall.



(c) Trafikpåverkan kan justeras samtidigt som stöd ges i form av vad AI-modellen tror att trafikpåverkan mest troligt är.

(d) Förklaring till prediktion.

Figur 7.1: Nuvarande, samt förslag på, visuell presentation och interaktion av resultatet från AI-modellen

Kapitel 8

Slutsatser och rekommendationer för framtida arbete

8.1 Trafikhändelser

Vår AI modell visar att det finns potential att förbättra prognoser av längden av trafikstörningar för händelser. Det största steg framåt kan förmodligen göras genom att förbättra kvalitén av inputdata. När Trafikverket började samla in data var det med ett annat syfte än att använda den som input för AI modeller. Det betyder att nuförtiden finns andra krav på datakvalité än just då. För att utvecklingen och produktionssättning av ett AI system går så smidigt som möjligt är det bra om man på förhand tänker på vilka krav systemet kommer att ha och designa en infrastruktur som uppfyller dem krav. Att anpassa ett legacy system på efterhand kan absolut funka, men är inte idealiskt. Baskontraktdata, till exempel, kom från PDF-filer som då överfördes till ett excel-ark med en mängd olika strukturer så att det var i stort sett omöjligt att extrahera 100% av informationen. Det tar mycket tid och information är inte alltid komplett eller korrekt. Så det är bättre att designa en bra data-arkitektur från början.

Utänför IT-infrastrukturen behöver man också ha tydliga verksamhetsinstruktioner och -regler på plats. Det finns en större variation i data om den är genererat av människor än när den är automatiskt genererad. Olika människor kommer alltid att jobba på lite olika sätt. Men man kan försöka att minska den variation genom att ha uppgifter och regler som är klara och tydliga. Speciellt när flera händelser inträffar samtidigt och operativ personal har mycket att göra samtidigt hjälper tydliga riktlinjer att strömlinjeforma processen. Vi hittade till exempel händelser som stängdes i systemet mer än en vecka efter att själva händelse avslutades. Då kan man ifrågakälla hur korrekt informationen fortfarande är.

Vid osäker data finns det flera andra saker man kan undersöka för att förbättra modellen. Kan man producera stora mängder syntetisk data för att förbättra modellen? Det är ett ganska nytt forskningsområde, och vi testade några saker med *svd* biblioteket, men det är värt att gräva djupare. Man kan också titta på att vikta färsk information tyngre än äldre information så att modellen bättre reflekterar nyare data och därmed aktuella arbetsmetodiker. Vi använde data från 2015 till och med 2021 för att träna vår modell, men kanske är det så att data från 2021 är en bättre reflektion av den nuvarande förhållanden och arbetssätt än data från 2015. Mycket kan hända på 6 år.

8.2 Tågs ankomst till station

AI-modellers användning vid prognostisering av tågs ankomst och avgång från och till stationer är av värde och fungerar med goda resultat. AI-prognoser är bättre än manuella prognoser, även vid direkt jämförelse (se Figur 4.12). Vidare är AI-genererade prognoser skalbara (se Figur 4.9) och kan utföras för samtliga platser i Sverige (se Figur 4.6b).

Det framtida arbetet för att sätta prognoser för tågs ankomst och avgång från respektive station är till stor del baserat på tillgång till realtidsdata och implementation av kontinuerliga prognoser i realtid. Att bygga upp systemet att fungera i realtid utvecklar inte bara systemets funktionalitet utan skapar även en ny dimension av features att tillgå i modelleringen som kretsar kring tid. Aspekter som inte tagits med i detta projekt är vad klockan är i varje trafiksituation och vart tåget borde vara vid denna tidpunkt. Detta är något som skulle möjliggöra flertalet vägar till

nya features samt en slags trigger till när prognoser behöver sättas. Om tidtabellen tyder på att tåget borde vara vid en specifik hållplats, men att så inte är fallet, så skulle man potentiellt sätt anse detta som en trigger till att börja generera prognoser för detta tåg. Detta för att tåget med största sannolikhet är försenat och kommer att behöva en prognos. Att kontinuerligt sätta prognoser i realtid för alla tåg i bruk kommer att behöva prioritering. I realtid kan man därför prioritera att sätta prognoser för de tåg som verkar behöva det.

Likt de flesta system som kretsar kring AI så gå det att optimera dess delar i all evighet. På grund av avgränsningar i projektet har AIRT varit tvugna att sätta feature freeze och avgränsa antalet modeller som testades. Utöver de två modeller som använts i Tåg i Tid, så är det av största intresse att se hur väl ett neuralt nätverk skulle hantera prognostiseringen. Mer specifikt är NN-metodiken LSTM (Long Short Term Memory) av intresse, som mer direkt hanterar sekvensiell tidsseriedata och vad som tidigare har hänt i tidsföljden, vilket är exakt det man eftersöker i denna lösning. Att använda sig av LSTM kräver dock en mer utvecklad datamodellering som kommer att kräva tid att bygga. Utöver neurala nät är också andra trädbaserade modeller av intresse. CatBoost ¹ är en trädbaserad modell som bättre hanterar kategoriska variabler än XGB och LGBM. Detta möjliggör att man istället för longitud och latitud kan beskriva platser i Sverige med platssignaturer.

Andra features att implementera i detta projekt finns det många av, dels kopplade till det enskilda tåget, men också kopplat till sträckan tåget befinner sig på, tågen längs denna sträcka och stationen vid vilken prognoser ska sättas. Samtliga features som AIRT har sett som möjliga förbättringar till nuvarande modellering återfinns i Tabell 8.1.

Tåget själv
Antal vagnar Händelse kopplat till tåget Hur lång tid tåget har kört från startdestination
Sträckan
Trafikstörande händelser, orsakskod, prognos trafikpåverkan Lutning/stigning Antal spår Historisk försening för samma tågtyp (tidigare avgång/dag/veckor/månad) Väder
Andra tåg på samma sträcka
Försening för tåg(en) som är framför på nuvarande sträcka (max/min/medel) Antal tåg framför Se "Tåget själv" (vid dynamiskt nätverk)
Station
Senast försening på samma spår som tåget ska in på (orsakat av stationen) Avgång – Ankomst i genomsnitt senaste timmen (för att se hur snabba tågen är därifrån) Avgång – Ankomst vid tidigare tillfällen av per tågtyp (dag/vecka) Genomsnittsförsening senaste kvart/timme per tågtyp Genomströmning av tåg vid samma tidpunkt (timme/dag/vecka/månad) Ankommande/avgående tåg (från andra sträckor), antal/typ Antal planerade spår (som är med i tidtabellen)
Kommande sträckor/stationer fram till slutdestination
Se "Sträckan" Påverkan av hastighetsnedsättningar längs sträckorna i minuter Förseningar som uppkommit under sträckan (under tiden som tåget förväntas befinna sig på sträckan) Trafikstörande händelser som är aktuella när tåget väntas vara där.

Tabell 8.1: Önskvärda features att implementera i framtida arbete.

¹<https://catboost.ai/>

8.3 Vädrets påverkan på järnvägssystemet

8.3.1 Framtida ”Järnvägsväder” och vädertyper

Först och främst är fortsatta studier av väder och händelser i järnvägssystemet önskvärda i syfte att utöka data-underlaget för att få säkrare resultat, vilket exempelvis beskrivs i 5.2.1. Utifrån de resultat som projektet fått fram avseende samband mellan väder och händelser i järnvägssystemet är det viktigt att ta vara på möjligheten att i framtiden predicera risken att händelser av en viss typ kan inträffa, baserat på en anpassad väderprognos. Det innebär framtagande och test av flera vädertyper och grunden i detta arbete utgörs dels av insamling av kunskap från berörda verksamheter som Trafikverket och järnvägsföretagen och dels en helhetssyn som även involverar meteorologisk kompetens.

Visionen i fokusområde väder har varit att den kunskap som byggs upp inom AIRT om samband mellan väder och händelser i järnvägssystemet i förlängningen ska kunna leda till prediktioner som gör järnvägstransporter mer effektiva och attraktiva. För att nå detta måste väderinformationen anpassas till järnvägens behov. Begreppet ”Järnvägsväder” introducerades som ett samlingsnamn för algoritmer och nya variabler som behöver identifieras för att beskriva vädrets konsekvenser för järnvägen. Detta utmynnade i ca 30 vädertyper som bygger på olika kombinationer av meteorologiska variabler men även en del nya variabler.

8.3.2 Kartläggning av järnvägsnätet

I 5.1.1 argumenterades det för hur viktigt det är att anpassa tillgängliga data till järnvägen och därvid inkludera ett meteorologiskt perspektiv. En första anpassning, baserad på höjd över hav, bör ses som ett första steg i en process där man eftersträvar en så hög anpassningsgrad som möjligt för att beskriva en plats eller delsträcka i järnvägsnätet. Exempel på data som bör ingå i framtida arbete är data som karakteriserar den närliggande terrängen, som bergstoppar och dalgångar, närhet till öppna vattenytor, markens beskaffenhet. Faktorer som kan ha betydelse för mikroklimatet är inte minst växtlighet och skuggor som exempelvis kan påverka yttemperaturen (eller rälstemperaturen) markant.

8.3.3 Introduktion av nya variabler

”Järnvägsväder” kan i framtiden leda till användning av avancerade branschpassade väderprognoser inom järnvägstransport som predicerar både väder- och markvariabler, d.v.s. väder och dess konsekvenser. Ytterligare ett steg i anpassningen är introduktion av variabler som beskriver händelser, som risk att händelser inträffar, hur stor risken är, osäkerheter mm. Om man kan predicera risken för en viss typ av händelser (orsakskod) kan en sådan mycket förädlad prognos utgöra bättre beslutsunderlag.

Framtida arbete borde fokusera på framtagande av flera vädertyper. Vädertyper är synonymt med väderkategorier bestående av variabler, kombination av variabler, och variabelernas värden, som är relevanta för järnvägssystemet och kan utgöra beslutsunderlag. I dagsläget använder Trafikverket relativt grova samband, och i vissa fall skulle dessa kunna omdefinieras. Detta kan illustreras med två exempel som skulle kunna föreslås för framtida studier. Det första är att man för bedömning av risk för solkurvor använder tröskelvärdet av variabeln lufttemperatur, som exempelvis 25 eller 30 °C, i stället för att ta fram en mer lämplig variabel som rälstemperatur eller yttemperatur. Det andra exemplet avser en vidareutveckling och anpassning av befintliga meteorologiska variabler till järnvägens behov med exemplet kraftig vind. Vindstyrkan kan anpassas till järnvägen genom att, utifrån banans riktning och exempelvis tågnummer beräkna motvind och sidvind, vilket skulle kunna bidra till ett bättre beslutsunderlag.

Både AI-relaterat arbete och fokus väder fann att det kan finnas samband mellan summan av nederbörden eller temperaturen i närtid, exempelvis tidsperioder på några dagar, och variabelerna händelsefrekvens respektive trafikpåverkans längd. Därför finns det en god anledning att definiera variabler som representerar summor av värden under en tidsperiod, exempelvis de senaste 3 dagarna, vilket framtida forskning bör ta vara på.

8.3.4 Geografiska skillnader

När det gäller fokus väder finns resultat som tyder på att det inom ett och desamma driftledningsområde kan förekomma stora skillnader i trafikpåverkans längd, som kan vara systematiska. Framtida forskning skulle dra nytta av att närmare undersöka dessa skillnader.

8.4 Framtidens visualisering genom digitalatvillingar

Med Digitala Tvillingar är målet att avbilda verkligheten digitalt som bekläs med information om aktuella lägesbild från verkligheten.

Digitala tvillingar kan användas till många andra exempel som skulle kunna vara: - Inspektion av bangården. - Evakuering och röjning, när tåg skall utrymmas. - Mikroklimat för att upptäcka solkurvor och trädbeväxta bandelar utifrån solinstålning.

Information om transport systemet kan användas i en förädlad form inom Digitala Tvillingar. Det ger oss möjlighet att identifiera komplexa samband som annars inte var möjliga att identifiera. Detta gäller både fordons transporter på väg och järnvägs transporter för både gods och person trafik.

Litteraturförteckning

- [1] M. Tiger, D. Jönsson, G. Eilertsen, A. Ynnerman, and F. Heintz, "Trafikverket - projekt ai," Linköpings Universitet, Tech. Rep., 2019. [Online]. Available: https://fudinfo.trafikverket.se/fudinfoexternwebb/Publikationer/Publikationer_004101_004200/Publikation_004135/LiU%20rapport%20-%20slutgiltig.pdf
- [2] "Forskningens utmaning kring utveckling av system med avancerad visualisering : Analys av barriärer och vägval," Tech. Rep. 2021:104, 2021, - AIRT-projektet (AI-baserad Realtidsprognostisering av Trafikinformation, TRV2020/5711) i samverkan med FOI-projektet "Styrning mot ökad punktlighet" (TRV 2020/26594) och "Digitaliserade beslutsstöd – policy och praktik" (TRV 2019/119686).
- [3] A. Baevski, Y. Zhou, A. Mohamed, and M. Auli, "wav2vec 2.0: A framework for self-supervised learning of speech representations," *Advances in Neural Information Processing Systems*, vol. 33, pp. 12 449–12 460, 2020.
- [4] J. Shen, R. Pang, R. J. Weiss, M. Schuster, N. Jaitly, Z. Yang, Z. Chen, Y. Zhang, Y. Wang, R. Skerrv-Ryan *et al.*, "Natural tts synthesis by conditioning wavenet on mel spectrogram predictions," in *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE, 2018, pp. 4779–4783.
- [5] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.
- [6] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A convnet for the 2020s," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 11 976–11 986.
- [7] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022, pp. 10 684–10 695.
- [8] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.
- [9] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell *et al.*, "Language models are few-shot learners," *Advances in neural information processing systems*, vol. 33, pp. 1877–1901, 2020.
- [10] NLLB Team, M. R. Costa-jussà, J. Cross, O. Çelebi, M. Elbayad, K. Heafield, K. Heffernan, E. Kalbassi, J. Lam, D. Licht, J. Maillard, A. Sun, S. Wang, G. Wenzek, A. Youngblood, B. Akula, L. Barrault, G. Mejjia-Gonzalez, P. Hansanti, J. Hoffman, S. Jarrett, K. R. Sadagopan, D. Rowe, S. Spruit, C. Tran, P. Andrews, N. F. Ayan, S. Bhosale, S. Edunov, A. Fan, C. Gao, V. Goswami, F. Guzmán, P. Koehn, A. Mourachko, C. Ropers, S. Saleem, H. Schwenk, and J. Wang, "No language left behind: Scaling human-centered machine translation," 2022.
- [11] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *International Conference on Knowledge Discovery and Data Mining (KDD)*, 2016.
- [12] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," *Advances in neural information processing systems*, vol. 30, 2017.
- [13] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "Catboost: unbiased boosting with categorical features," *Advances in neural information processing systems*, vol. 31, 2018.

- [14] C. Bentéjac, A. Csörgő, and G. Martínez-Muñoz, “A comparative analysis of gradient boosting algorithms,” *Artificial Intelligence Review*, vol. 54, no. 3, pp. 1937–1967, 2021.
- [15] I. Shavitt and E. Segal, “Regularization learning networks: Deep learning for tabular datasets,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.
- [16] L. Katzir, G. Elidan, and R. El-Yaniv, “Net-dnf: Effective deep modeling of tabular data,” in *International Conference on Learning Representations (ICLR)*, 2020.
- [17] A. Kadra, M. Lindauer, F. Hutter, and J. Grabocka, “Regularization is all you need: Simple neural nets can excel on tabular data,” *arXiv preprint arXiv:2106.11189*, 2021.
- [18] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. De Freitas, “Taking the human out of the loop: A review of bayesian optimization,” *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148–175, 2015.
- [19] J. Bergstra, D. Yamins, and D. Cox, “Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures,” in *International conference on machine learning*. PMLR, 2013, pp. 115–123.
- [20] D. Engelsons, M. Tiger, and F. Heintz, “Coverage path planning in large-scale multi-floor urban environments with applications to autonomous road sweeping,” in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 3328–3334.
- [21] A. Paleyes, R.-G. Urma, and N. D. Lawrence, “Challenges in deploying machine learning: a survey of case studies,” *ACM Computing Surveys (CSUR)*, 2020.
- [22] B. Klievink, B.-J. Romijn, S. Cunningham, and H. de Bruijn, “Big data in the public sector: Uncertainties and readiness,” *Information systems frontiers*, vol. 19, no. 2, pp. 267–283, 2017.
- [23] L. A. Castelijns, Y. Maas, and J. Vanschoren, “The abc of data: A classifying framework for data readiness,” in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 2019, pp. 3–16.
- [24] N. D. Lawrence, “Data Readiness Levels,” *ArXiv e-prints*, May 2017.
- [25] Z.-C. Li and D. Sheng, “Forecasting passenger travel demand for air and high-speed rail integration service: A case study of beijing-guangzhou corridor, china,” *Transportation Research Part A: Policy and Practice*, vol. 94, pp. 397–410, 2016.
- [26] M. Milenković, L. Švadlenka, V. Melichar, N. Bojović, and Z. Avramović, “Sarima modelling approach for railway passenger flow forecasting,” *Transport*, vol. 33, no. 5, pp. 1113–1120, 2018.
- [27] Z. Fang, Q. Cheng, R. Jia, and Z. Liu, “Urban rail transit demand analysis and prediction: A review of recent studies,” in *International Conference on Intelligent Interactive Multimedia Systems and Services*. Springer, 2018, pp. 300–309.
- [28] T. Xia, X. Song, Z. Fan, H. Kanasugi, Q. Chen, R. Jiang, and R. Shibasaki, “Deeprailway: A deep learning system for forecasting railway traffic,” in *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*. IEEE, 2018, pp. 51–56.
- [29] P. Kecman and R. M. Goverde, “An online railway traffic prediction model,” in *RailCopenhagen2013: 5th International Conference on Railway Operations Modelling and Analysis, Copenhagen, Denmark, 13-15 May 2013*. International Association of Railway Operations Research (IAROR), 2013.
- [30] —, “Predictive modelling of running and dwell times in railway traffic,” *Public Transport*, vol. 7, no. 3, pp. 295–319, 2015.
- [31] H. Park and A. Haghani, “Real-time prediction of secondary incident occurrences using vehicle probe data,” *Transportation Research Part C: Emerging Technologies*, vol. 70, pp. 69 – 85, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0968090X15000972>

- [32] S. Alexandra, "Orsakskoder rapport," in *Orsakskoder A-Ö*, 18th ed., E. Arvidsson, Ed. Trafikverket, 2018, p. 41. [Online]. Available: <https://trvdokument.trafikverket.se/fileHandler.ashx?typ=showdokument&id=e258396e-4078-43d7-b256-e8513f974806>
- [33] D. W. S. Donald E. Hilt, *Ridge, a computer program for calculating ridge regression estimates*, 1977. [Online]. Available: <https://doi.org/10.5962/bhl.title.68934>
- [34] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.
- [35] T. K. Ho, "Random decision forests," *Proceedings of 3rd international conference on document analysis and recognition*, vol. 1, pp. 278–282, 1995.
- [36] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of Statistics*, p. 1189–1232, 2001.
- [37] R. E. Schapire, "Explaining adaboost," *Empirical inference*, pp. 37–52, 2013.
- [38] W. P. W. S. McCulloch, "A logical calculus of the ideas immanent in nervous activity," *The Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115–133, 1943.