Triple F Rapport nummer: 2022.4.2.2

2023

Självlärande neurala nätverk för operativ lokstyrning – SOL förstudie

SARA GESTRELIUS, ZOHREH RANJBAR, RISE

SLUTRAPPORT

TRIPLEF
FOSSIL FREE FREIGHT

Projektnummer 2022.4.2.2	
Titel på projektet – svenska Självlärande neurala nätverk för operativ lokstyrning – SOL förstudie	
Titel på projektet – engelska Self-learning neural network for operational locomotive control - SOL prestudy	
Projektledareorganisation RISE	
Namn på projektledare Zohreh Ranjbar	
Namn på ev övriga projektdeltagare Sara Gestrelius Jonatan Gjerdrum Henrik Wester	
Nyckelord: 5-7 st Machine learning, Reinformnet learning, deep learning, optimisation, operativ styrning	

Sammanfattning

Förstudien SOL fokuserar på trafikeffektivitet, resursoptimering och AI-baserat beslutstöd för godstågoperatörer, vilket bidrar till Triple Fs mål om ett mer transporteffektivt samhälle och minskade CO2-utsläpp. Bra beslutsstöd för den operativa planering av godståg kan ge en förbättrad resursallokering och störningshantering, vilket i sin tur har en positiv påverkan på transportkostnader och pålitlighet. Minskade kostnader och bättre pålitlighet ökar järnvägens konkurrenskraft och kan således också bidra till överflyttning av gods från väg till järnväg.

Mer specifikt handlar förstudien om hur självlärande algoritmer med djupa neurala nätverk (DRL) skulle kunna stötta lokplanerare. Dataanalys visar att man på Green Cargo jobbar aktivt med lokstyrning, och loken bokas frekvent om. Det är ofta svårt att fullt ut förstå hur en ombokning av ett lok påverkar trafiken och transportkvaliteten, och därför skulle ett intelligent beslutsstöd vara av nytta. Ett av resultaten från SOL är en beskrivning av hur en AI-baserad agent kan utvecklas och tränas för att stötta en lokplanerare, men det krävs ytterligare forskning för att se om agenten kan fungera som ett adekvat beslutsstöd. Projektet har bidragit till kunskapsuppbyggnad inom järnvägssektorn och har potential att minska kostnader, förseningar och förbättra kundkommunikationen.

Nästa steg innefattar större genomförande- eller doktorandprojekt för vidare studier av den föreslagna metoden. Förutom en mer omfattande fullstudie av hur en AI-agent kan tränas för att stötta en lokplanerare, inkluderar framtida utmaningar även datatillgänglighet, implementering och att göra AI användbart och förståbart för järnvägssektorn.

Summary

The SOL feasibility study focuses on traffic efficiency, resource optimization and AI-based decision support for freight train operators, which contributes to Triple F's goals of a more transport-efficient society and reduced CO2 emissions. Good decision support for operational planning of freight trains can improve resource allocation and disturbance management, which in turn have a positive impact on transport costs and reliability. Reduced costs and increased reliability increase the competitiveness of rail and can thus contribute to shifting goods from road to rail.

More specifically, SOL considers how Deep Reinforcement Learning (DRL) could support locomotive planners. Data analysis shows that planners at Green Cargo, a large freight railway undertaking in Sweden, frequently change the locomotive plan during operations to handle various disturbances and/or to optimize the plan. It is often difficult to fully understand how rerouting a locomotive affects traffic and transport quality, and therefore intelligent decision support tools would be beneficial. One of the results from SOL is a description of how an AI-based agent can be developed and trained to support a locomotive planner, but further research is needed to see if the agent can provide adequate decision support. The project has contributed to knowledge building on AI within the railway sector and has the potential to reduce costs, reduce delays and improve customer communication.

The next steps include larger implementation or doctoral projects to further develop and investigate the proposed method. In addition to a more comprehensive study of how an AI agent can be trained to support a locomotive planner, future challenges include data availability, implementation and making AI usable and understandable for the railway sector.

Innehållsförteckning

Sammanfattning **2**

Summary **2**

Innehållsförteckning **3**

1 Inledning **4**

2 Bidrag till Triple F **4**

3 Genomförande **5**

4 Resultat **6**

4.1 Operativ planering **6**

4.2 Datainsamling, analys och kvalitet **8**

4.3 Föreslagen metod och initial utvärdering **10**

4.3.1 DRL-Agenten 11

4.3.2 DRL-träning genom simulering 14

4.3.3 Ett första försök 15

5 Nyttiggörande och nästa steg **17**

6 Diskussion **18**

Referenslista **19**

1 Inledning

Förstudien SOL genomfördes i samarbete mellan RISE och Green Cargo för att undersöka möjligheten och nyttan av att använda artificiell intelligens (AI) för att stödja den operativa lokplaneringen. Det finns en brist av stöd för optimerad planering inom järnvägsbraschen, vilket är olyckligt då bättre operativ planering skulle kunna hjälpa branschen med två av dess största utmaningar: kostnader och tillförlitlighet.

Om järnvägstransporterna blir billigare och/eller mer pålitliga blir de också mer attraktiva för transportköpare. I en ”stated preference” undersökning från början av 2000-talet där olika faktorer vid val av transport rangordnades var de tre viktigaste faktorerna ”kostnadsminskningar på 5%”, ”kostnadsminskningar på 3%” och ”förseningsminskning på 30%” (Lundström, 2006). Ett annat resultat i samma rapport är att ca 80% av transportköparna skulle byta till en likvärdig transportör om det gav en kostnadsminskning på 5% eller mer. I en workshop med transportköpare från 2014 återfinns liknande resultat – det som skulle öka intresset för järnvägstransporter mest är en prissänkning (Nelldal & Boysen, 2014). I (Nelldal & Boysen, 2014) påpekas även att järnvägen har dåligt rykte när det kommer till pris, godshantering och tidhållning. Tidhållningen handlar i det här fallet inte om minuter, utan om dagar. Transportköparna gav också uttryck för att åkerierna är ”tydligare och säkrare i sin kommunikation” vilket tillskrevs det faktum att åkerierna har bättre kontroll på sin verksamhet, och inte har samma problem med att få tillgång till infrastruktur (väg/spår) som järnvägsoperatörerna. Sist men inte minst pekar (Flodén, Bärthel, & Sorkina, 2017) åt samma håll. De identifierar kostnad, transporttid, pålitlighet och kvalitet som viktiga faktorer för val av transportmedel, givet att transporterna uppfyller vissa baskrav. Ett beslutstöd som hjälper godsoperatörerna att minska kostnaderna, minska förseningar, öka tillförlitlighet och få bättre kontroll över sina resurser borde således leda till att järnvägstransporter blir mer konkurrenskraftiga, vilket i sin tur kan leda till att gods flyttas från väg till järnväg.

Green Cargo, Sveriges största godstågoperatör, står dagligen inför utmaningen att planera och upprätthålla lokomlopp. Komplexiteten i utmaningen ges inte bara av lokomloppen utan även av andra resurser såsom förare, vagnar och tåglägen. Utöver att skapa en bättre operativ plan skulle Green Cargo också vilja förbättra sin kommunikation med kunder vid försenade leveranser samt bli bättre på att överföra den erfarenhet som är avgörande för att fatta bra operativa beslut.

Hypotesen i förstudien är att självlärande algoritmer med djupa neurala nätverk (deep reinforcement learning, DRL) kan stötta lokledare genom att föreslå omloppsändringar, och på så sätt förbättra den operativa lokplaneringen. Tanken är att DRL-agenten ska kunna lära sig att generera ändringsförslag och förutspå tågankomster genom neurala nätverk och passande belöningssystem. Detta tillvägagångssätt bygger på den senaste utvecklingen inom maskininlärning och optimering, och tidigare framgångsrika DRL-projekt visar potential för att överträffa människoskapade metoder inom olika områden. På lång sikt förväntas projektet bidra till bättre beslutsfattande, optimerad omplanering av resurser och förbättrad kommunikation med kunder vid försenade leveranser.

Projektet startades oktober 2022 och avslutades juli 2023. Finansieringen har erhållits från Trafikverket inom ramen för programmet Triple F, samt i form av inkind från Green Cargo.

2 Bidrag till Triple F

För att åstadkomma en överflyttning av inlandstransporter från väg till det mer energieffektiva transportslaget järnväg krävs att järnvägstransporter blir mer konkurrenskraftiga. Två komponenter som påverkar järnvägens attraktivitet är pålitlighet och pris. Båda dessa komponenter påverkas i sin tur bland

annat av hur pass bra godstågsoperatörerna är på att allokera och styra sina resurser, dvs. hur pass bra man planerar – och vid störningar omplanerar – hur lok, förare och vagnar ska användas.

Förstudien fokuserar på lokplanering, men som förklaras längre ner i rapporten kan kunskap och teknik för lokplanering till viss del också överföras till andra resursslag. På kort sikt kommer projektet att bidra till kunskapsuppbyggnad och stötta utvecklingen av datadrivna beslutsstödssystem inom järnvägssektorn. Den kunskap som projektet bygger upp är relevant för järnvägsoperatörer, både i Sverige och internationellt. Det övergripande målet är en järnvägsindustri med effektivare resursstyrning och mer välunderbyggda beslut. På lång sikt förväntas projektet kunna minska kostnader, minska förseningar och förbättra kundkommunikationen. Dessa faktorer kan öka järnvägens konkurrenskraft gentemot vägtransporter och leda till överflyttning av gods till järnvägen.

Projektet SOL har således en tydlig koppling till Triple Fs programmål. Genom att fokusera på trafikeffektivitet, resursoptimering och AI-baserat beslutstöd för godstågoperatörer bidrar projektet till målet om ett mer transporteffektivt samhälle. Genom att främja övergången från fossila bränslen till fossilfria alternativ inom godstransportsystemet stödjer projektet även målet om reducering av CO2-utsläpp. Över 95 procent av Green Cargos transportarbete sker med eltåg vilket ger ett mycket lågt CO_2 utsläpp (Green Cargo, 2014).

3 Genomförande

Förstudiet SOL med följande arbetspaket och aktiviteter har genomförts:

WP1: Dataanalys och kunskapsinsamling för lokomlopp

Förutom flertalet möten inom projektgruppen för att samla in kunskap och diskutera modellering, datatillgång och dataanalys så hölls ett antal workshoppar/intervjuer med externa personer. Data inhämtades och analyserades.

Resultat från WP1 presenteras i 4.1 och 4.2.

WP2: Träningsmiljö för lokomlopp

En träningsmiljö bestående av en enkel simulering av MB-lokomloppen mellan Borläge och Luleå utvecklades. Vidare diskuterades hur en träningsmiljö skulle kunna se ut i en storskalig studie.

Resultat från WP2 presenteras i 4.3.2 och 4.3.3.

WP3: DRLA-utveckling och teknisk utvärdering

DRL-agentens tillstånd, handling samt belöning modellerades och implementerades i Python. Initiala försök med att träna agenten genomfördes och analyserades.

Resultat från WP3 presenteras i 4.3.1 och 4.3.3.

WP4: Rapportering och genomförandeprojekt

Under hela projektet har möten för statusrapportering och uppföljning genomförts med alla i projektteamet. En slutpresentation hölls också för relevant personal på Green Cargo. Under denna presentation lyftes frågeställningen om det skulle vara intressant med ett fullskaligt genomförandeprojekt.

Projektet har producerat en delrapport, samt denna slutrapport.

4 Resultat

4.1 Operativ planering

I det här avsnittet beskrivs kort aspekter som är viktiga att beakta vid operativ lokstyrning, samt hur dessa relaterar till DRL. Materialet om lokstyrning kommer från workshoppar och intervjuer. Dessa listas i Tabell 1. Initialt diskuterades alla aspekter som kan tänkas påverka lokomloppen, men vid senare workshoppar avgränsades frågeställningen till lok, lokomlopp, verkstäder, underhåll samt trafik.

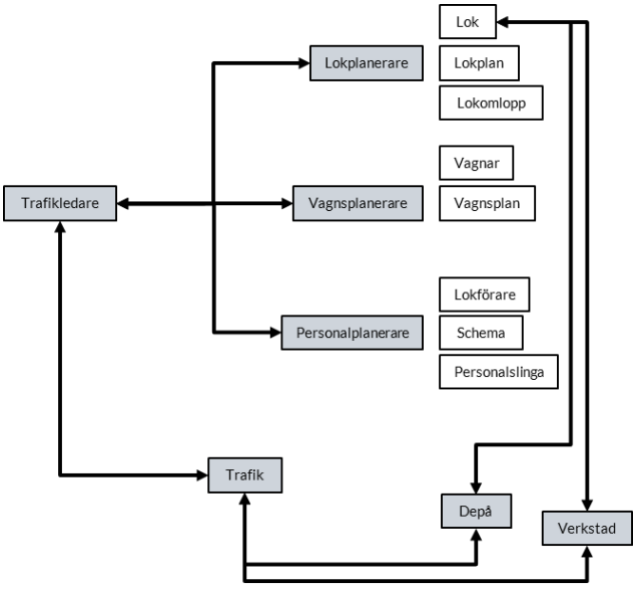
Tabell 1: Intervjuer och workshoppar.

Datum	Möte	Närvarande
11 november 2022	Inledande workshop med personer som tidigare forskat på omloppsplanering för godståg.	Zohreh Ranjbar, Sara Gestrelius, Martin Aronsson, Martin Joborn.
21 november 2022	Workshop 1 för att rita upp hur den operativa planeringen går till.	Projektteamet + personal från Green Cargo.
22 november 2022	Workshop 2 för att rita klart processkartan för operativ planering.	Projektteamet + personal från Green Cargo.
2 december 2022	Möte för att identifiera datakällor samt initiala diskussioner om vad som är bra respektive dålig lokstyrning.	Projektteamet + personal från Green Cargo.
20 januari 2023	Möte för att diskutera databehov och källor.	Projektteamet + personal från Green Cargo.

Under de två workshopparna i november med personal från Green Cargo ritades ett detaljerat diagram över den operativa resursstyrningen upp. Utifrån detta detaljerade diagram har vi tagit fram ett

sammanfattande och mer generiskt diagram med fokus på lokstyrning (Figur 1). Ett lokomlopp är ett antal tåg som kan köras med samma lokindivid. Ett lokomlopp består således av ett antal tåg och kopplingar, där en koppling från tåg A till tåg B innebär att loket ska dra tåg B direkt efter tåg A. En lokplan specificerar vilket omlopp som ett specifikt lok ska köra och när. En viktig aspekt av lokplanen är att loket med jämna mellanrum behöver köras till verkstaden för olika sorters tillsyn och/eller underhåll. Det finns gränser för hur långt ett lok får köra innan tillsyn /underhåll måste genomföras, och olika tillsyn/underhåll har olika gränser. Detta innebär att det är viktigt att se till att lok kan spendera tillräckligt lång tid på verkstäder tillräckligt ofta. Utöver verkstäderna och deras placering är även depåer viktiga för lokplanering. Depåer är spår där loken står parkerade när de inte drar tåg.

SOL förstudiens fokus är lokplanering. Det är dock värt att notera att omplanering för vagnar och personal på många sätt liknar lokplanering. De vita boxarna för varje resursslag har likheter, och alla resurser påverkas av trafiken (om ett tåg är sent innebär det att lok, vagnar och lokförare är sena). Faktum är att dessa tre olika planeringsproblem är så pass lika att om en DRL-agent tränas upp för lokplanering bör man kunna använda "transfer learning" för att lära upp separata agenter för personalplanering och vagnsplanering. Transfer learning innebär att man använder kunskap som erhållits från en annan modell eller uppgift och överför den till en ny modell eller uppgift. Genom att överföra kunskap från en tränad modell till flera agenter kan man snabba upp och förbättra träningsprocessen. Transfer learning kan användas för att förse agenterna med en gemensam grund av kunskap som de olika agenterna sedan kan bygga vidare på genom ytterligare träning på specifika uppgifter. I vårt fall är till exempel trafiken, dvs. hur tågen kör och förseningar propageras, samma för alla resursslag. Ytterligare en teknik som är användbar för operativ planering är ”orchestration”. En ”orchestrator” är ett program som styr samarbetet mellan flera DRL-agenter för att optimera prestanda och samarbete mot ett gemensamt mål. Vi kan således träna en agent per transportsslag, och låta en orchestrator hantera samarbetet. Detta liknar på många sätt dagens arbetssätt där trafikledaren/den operativa chefen har ett samordnande ansvar. Detta tillvägagångssätt möjliggör effektiv användning av tillgängliga resurser och öppnar upp möjligheter för mer avancerad och optimerad planering av både personal och vagnar i framtiden.



Figur 1: En grov skiss över operativ resursstyrning (med fokus på lokresurserna) på ett godstågbolag. Gråa rutor är personer eller artefakter som påverkar resursers styrning och/eller tillgänglighet. Vita rutor är resurser och olika sorters planer. Lokförare är i det här diagrammet endast ”resursen” lokförare.

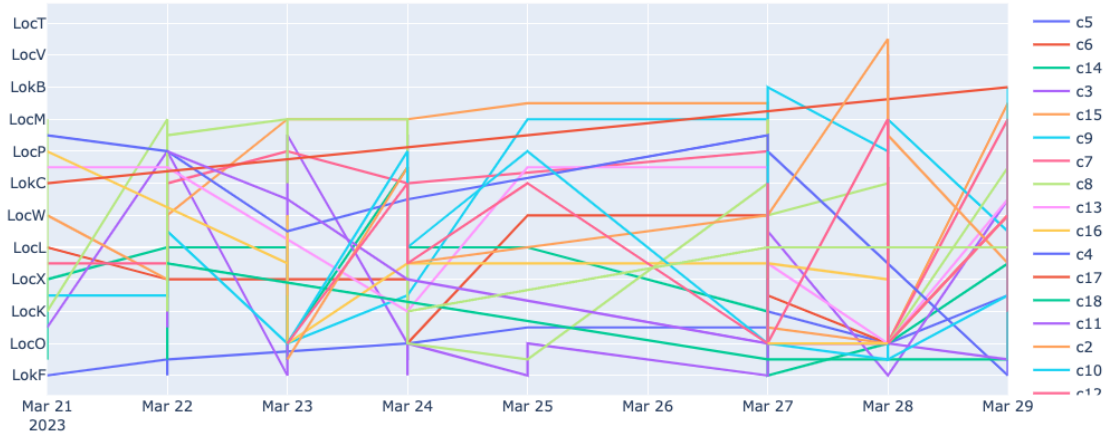
4.2 Datainsamling, analys och kvalitet

Förstudiens exempelfall är ståltågen på sträckan Borlänge-Luleå. Data hämtades från olika system, sammanställdes och analyserades. Målet med dataanalysen var att utvärdera tillgängliga datakällor, datakvalitet, förstå hur ofta lok planeras om samt hur ofta lok befinner sig på verkstad för underhåll. Analysen diskuterades med relevanta personer på Green Cargo, samt presenterades vid slutpresentationen.

Som ett provskott fick vi 9 dagars loggningsdata för lokomlopp på sträckan Borlänge-Luleå, från 21 mars 2023 till 29 mars 2023. I logg-filerna finns data från både innan och efter loggningsdatumet. Så i en logg-filen från till exempel 22 mars kunde vi således se både vilka lok som kört vilka tåg under föregående dagar, samt vilka lok som var planerade att köra vilka tåg i framtiden. På så sätt kunde vi följa när lok planerades om. Det saknades information om tågförseningar på grund av tekniska problem som uppstått när Trafikverket infört ett nytt planeringssystem. Detta begränsade möjligheten att se konsekvenserna av lokomloppsändringar i form av tågförseningar.

Lokomloppsanalys

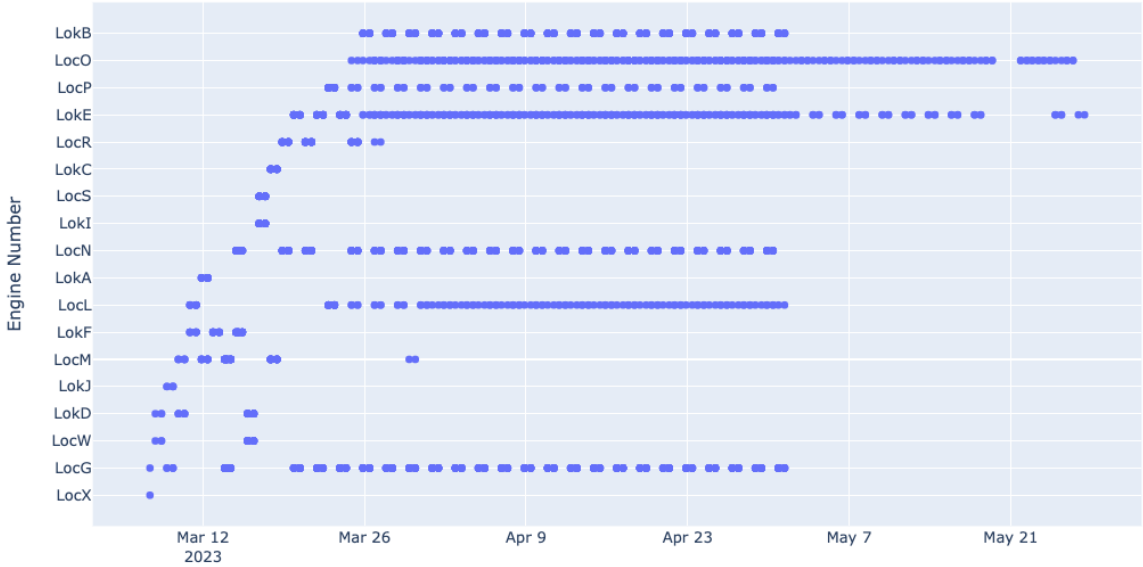
Diagrammet i Figur 2 visar hur MB-loken flyttas mellan olika lokomlopp. X-axeln är dagar, y-axeln lok och de färgade strecken är lokomlopp. Om ett lokomlopp hade haft samma lokindivid under hela perioden skulle detta vara ett rakt, horisontellt streck. Lokomlopp som körs av olika lok ser däremot ut som kantiga streck där varje ”böj” innebär att loket som kör omloppet bytts ut. Som synes ändrades lokomloppen betydligt under de nio dagarna.



Figur 2: Hur lok byts mellan olika lokomlopp.

Lokplanering för ett enskilt tåguppdrag

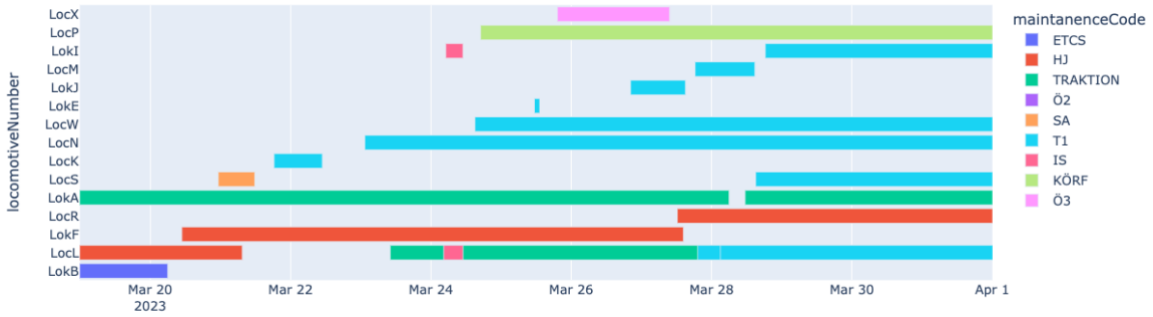
Diagrammet i Figur 3 illustrerar lokplaneringen för ett enskilt tåguppdrag under olika dagar mellan mars och maj. En punkt innebär att ett lok varit planerat att köra tåget enligt loggningsdatafilerna. För tidpunkter som ligger innan den 21 mars innebär detta att punkterna visar vilket lok som faktiskt kört tåget. För punkter mellan 21 mars och 29 mars finns punkter för de lok som faktiskt kört tåget, samt punkter för lok som varit planerade att köra tåget men sedan inte kört det. För tåg efter 29 mars finns punkter för lok som någon gång mellan 21 och 29 mars varit planerade att köra tåget. Notera att ståltåg generellt kräver två lok och därför finns det två punkter för de första dagarnas tåg. Som synes verkar långtidsplanen, dvs. de lok som är planerade att köra tåget i april och maj, inte ha någon större betydelse för vilket lok som i slutändan kör tåget. Detta beror på att det var så pass vanligt med ändringar i lokallokeringen under mars månad att den operativa planeringen inte alls följde den ursprungliga långtidsplanen.



Figur 3: Lokplanering för ett enskilt tåguppdrag.

Verkstadsdata

Analys av verkstadsdata visar att det är vanligt att flera lok befinner sig på verkstad samtidigt. Anledningen till verkstadsbesöken varierar dock, och innehåller allt från översyn till avhjälpande underhåll till investeringsåtgärder. I diagrammet i Figur 4 är x-axeln tid och y-axeln lok. En horisontell stapel innebär att loket befann sig på verkstad under den tiden, och färgen på stapeln visar huvudorsaken till verkstadsbesöket. Om en stapel når hela vägen till 1 april betyder det att ingen sluttid var satt när vi fick data. Sluttiden sätts generellt inte förrän när man konstaterat att loket är redo att tas i drift igen, dvs. verkstadstiderna ser evighetslånga ut fram tills loken ”plötsligt” blir klara och är redo att användas. Vissa av staplarna som sträcker sig hela vägen till 1 april kan således i slutändan blir kortare. Att lok är på verkstad så ofta försvårar den operativa planeringen. Dels eftersom det innebär att färre lok är tillgängliga, dels eftersom det innebär att lok måste dras eller köras till verkstad.



Figur 4: Lok på verkstad.

Slutsatser

Att lokplaneringen ändras så mycket bekräftar behovet av beslutsstödsystem. Inte minst eftersom det kan vara svårt för en planerare att fullt ut uppskatta effekterna av en omplanering då den har påverkan i flera steg och på flera resursslag. Ett beslutsstödssystem kan hjälpa planeraren att hantera komplexiteten och bidra till att färre ändringar behöver göras. Beslutsstödet bör även underlätta optimering av planen och på så sätt möjliggöra en mer effektiv hantering av resurser och därigenom förbättrad driftsprestanda och minskade förseningar inom godstransportsystemet.

Vidare kan projektet fastställa att det på Green Cargo finns data av tillräckligt god kvalité för att kunna genomföra djupare analyser och utvärdera konsekvenserna av lokomloppsändringar. Förståelsen för verkstadsdata och behov av översyn och underhåll för lok är speciellt viktigt för att kunna förbättra effektiviteten och minska antalet verkstadsbesök.

4.3 Föreslagen metod och initial utvärdering

En viktig uppgift under förstudien är att undersöka hur ett AI-baserat beslutsstödsystem skulle kunna fungera och utvecklas. DRL-agenter lär sig genom neurala nätverk att kartlägga olika tillstånd i en miljö, samt vilka handlingar som är fördelaktiga genom ett belöningssystem (Arulkumaran, Deisenroth, Brundage, & Bharath, 2017). För att utveckla en DLR-agent krävs således att vi definierar tillstånd, handling och belöning. Resultatet från detta arbete beskrivs i ”4.3.1 DRL-Agenten”. Vidare behöver en strategi för träning tas fram, vilket beskrivs i stycke ”4.3.2 DRL-träning genom simulering”. Eftersom SOL är en förstudie och inte ett fullskaligt FoI-projekt diskuteras också hur de olika aspekterna kan

hanteras i ett fullskaligt projekt. Projektet har dessutom gjort en provkörning med ett enklare ramverk, och resultaten från detta presenteras i stycke ”4.3.3 Ett första försök”.

4.3.1 DRL-Agenten

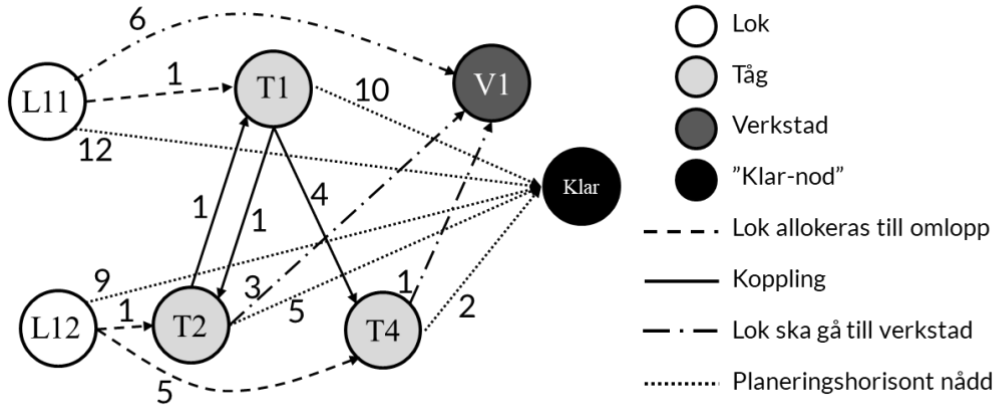
Val av algoritm

Efter att ha utforskat olika metoder och algoritmer för träning av agenten, kom vi fram till att PPO (Proximal Policy Optimization) är ett bra val för ändamålet. PPO är en av de mest använda algoritmerna inom reinforcement learning och har visat sig vara effektiv för att träna agenter att utföra komplexa uppgifter. Vi valde PPO eftersom det är en policy gradient-metod som fokuserar på att optimera en policy, vilket är det vi behöver för att lära vår agent att fatta beslut baserat på observationer och maximera den förväntade belöningen. Algoritmen använder en kombination av policy- och värderingsdelar för att generera handlingar och bedöma deras värde.

En av fördelarna vi ser med PPO är dess stegvisa optimering. Istället för att göra stora uppdateringar i en enda iteration, uppdaterar PPO policyn i små steg. Detta bidrar till en mer stabil och jämn inlärning av agenten, vilket är viktigt för att undvika överanpassning eller snabba förändringar i beteendet. Snabba förändringar i beteendet kan vara svåra för en planerare att hantera operativt.

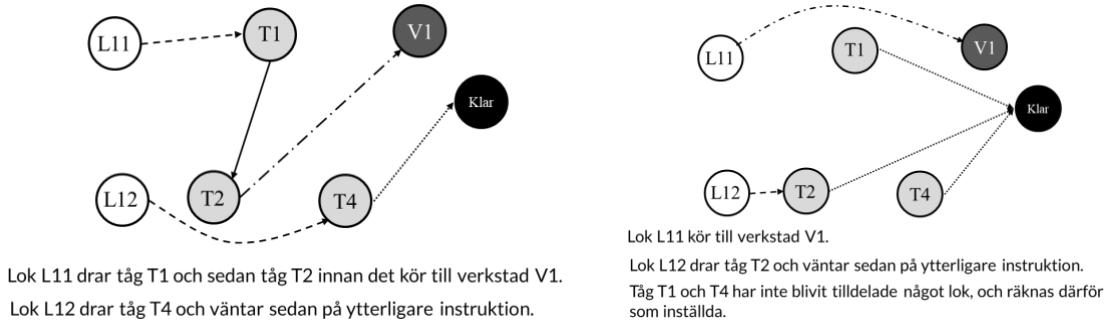
Lokplanering modellerat som en graf

Lokplanering kan modelleras som en graf där noder är (1) lok, (2) tåg, (3) verkstäder och (4) dummy-noder som markerar slutet på planeringshorisonten. Kanter representerar (1) möjliga kopplingar, (2) vilka tåg som kan vara ett loks första tåg, samt (3) vad som kan vara slutet på en planering. Planeringen kan antingen sluta med att planeringshorisonten nås, eller med att loket körs till en verkstad eller en depå. Kanterna har vikter som representerar kantens kostnad (dessa vikter kommer sättas av agenten). En exempelgraf med två lok, 3 tåg och en verkstad visas i ”Figur 5: Lokplaneringsproblemet som en graf.”. Kom ihåg att en koppling från tåg A till tåg B innebär att loket som drar tåg A sedan ska dra tåg B utan att dra något annat tåg däremellan.



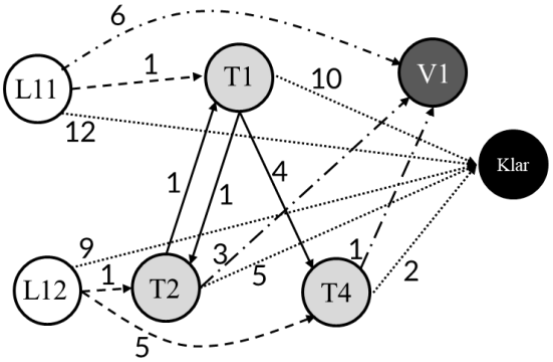
Figur 5: Lokplaneringsproblemet som en graf.

För att gå från en graf till lokplaner (och inställda tåg) kommer vi hitta icke-överlappande vägar som täcker alla lok- och tågnoder och som tillsammans kostar så lite som möjligt. En väg måste börja på ett lok eller ett tåg och måste sluta på en verkstad eller klar-noden. Om en väg börjar på ett lok så ska tågen i vägen dras av det loket. Om en väg börjar på ett tåg är alla tåg på vägen inställda (eftersom de inte tilldelats något lok). ”Figur 6: Två exempel på lokplaner.” visar två olika exempel.



Figur 6: Två exempel på lokplaner.

Problemet med att hitta de vägar som täcker alla lok och tåg till minsta kostnad kan lätt göras om till ett tilldelningsproblem som det finns effektiva algoritmer för att lösa. Tilldelningsproblemet består i att man ska tilldela arbetare till uppgifter. Olika arbetare tar olika mycket betalt för att utföra olika uppgifter, men alla uppgifter måste utföras och en arbetare kan göra max en uppgift. I vårt exempel är lok och tåg uppgifter, och tåg, verkstäder och sluthorisonten arbetare. Vi behöver se till att vi har lika många ”sluthorisont”-arbetare som den noden (”klar”-noden) har in-kanter, samt att vi har lika många ”verkstadsarbetare” som verkstaden har kapacitet för (eller lika många som den har in-kanter om verkstaden anses ha oändlig kapacitet). Kostanden för att tilldela en arbetare (tåg/verkstad/sluthorisont) till en uppgift (lok/tåg) är samma som kostnaden på den kant som går från uppgiften till arbetaren, eller om det inte finns någon sådan kant, en stor kostnad M. M måste vara tillräckligt stor för att aldrig ingå i den optimala lösningen, dvs. den måste dominera alla ”riktiga” kostnader. I ”Figur 7: Matris för tilldelningsproblemet.” syns den matris som specificerar tilldelningsproblemet för vår exempelgraf. Denna matris är allt som krävs för att kunna lösa tilldelningsproblemet, till exempel genom Kuhn-Munkres algoritmen, eller använda något av de metoder som finns i färdiga bibliotek, till exempel linear_sum_assignment-metoden från Optimize-modulen i SciPy (SciPy documentation, 2023).



$$
\begin{array}{c}
 \\ T1 \\ T2 \\ T4 \\ V1_{L11} \\ V1_{T2} \\ V1_{T4} \\ K_{L11} \\ K_{L12} \\ K_{T1} \\ K_{T2} \\ K_{T3}
\end{array}
\begin{array}{c}
\begin{array}{ccccc} L11 & L12 & T1 & T2 & T4 \end{array} \\
\begin{pmatrix}
1 & M & M & 1 & M \\
M & 1 & 1 & M & M \\
M & 5 & 4 & M & M \\
6 & M & M & M & M \\
M & M & M & 3 & M \\
M & M & M & M & 1 \\
12 & M & M & M & M \\
M & 9 & M & M & M \\
M & M & 10 & M & M \\
M & M & M & 5 & M \\
M & M & M & M & 2
\end{pmatrix}
\end{array}
$$

Figur 7: Matris för tilldelningsproblemet.

Tillstånd, handling och belöning

En DRL-agent lär sig genom att i olika *tillstånd* göra olika *handlingar* och notera vilken *belöning* (eller straff) detta genererar. Således behöver dessa tre aspekter definieras för agenten.

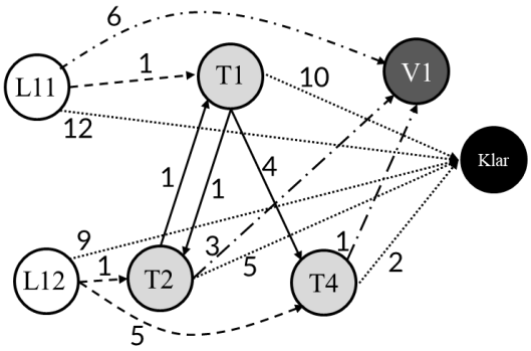
I Tabell 2 syns vilka data som kan ingå i en enklare modell av ett tillstånd. En aktuell koppling är en koppling som just nu är planerad att utföras. En giltig koppling är en koppling som är möjlig att genomföra. Notera att kopplingar kan gå från att vara giltiga till att bli ogiltiga, till exempel om det andra tåget i kopplingen avgår medan det första fortfarande inte har fått något lok. Ett annat viktigt koncept är buffertid. Buffertid är hur mycket extra tid som finns i en koppling. Säg att det tar 15 minuter att flytta ett lok från ankommande tåg T1 till avgående tåg T2. Om ankomsttiden för T1 är 09:00, och avgångstiden för T2 är 09:17, så blir buffertiden för kopplingen från T1 till T2 två minuter.

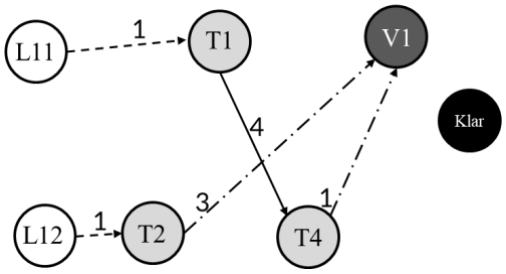
Tabellen specificerar en enklare modell av ett tillstånd där all data som krävs redan är tillgänglig och kan hanteras i en simulering. I en mer fullskalig modell bör allting som man vill att agenten ska beakta ingå, t.ex. även väder och trafikstatus.

Tabell 2: Bastillstånd.

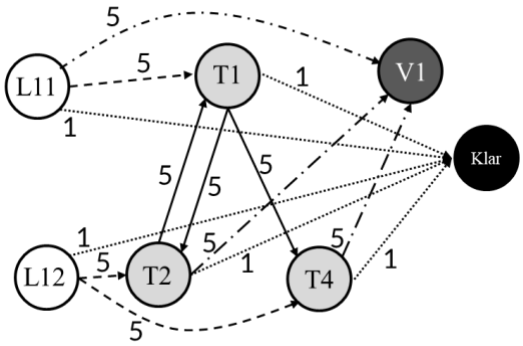
För lok (-noder):	Ort som loket befinner sig på, inspektions/underhållsgränser som måste uppfyllas (km), distans sedan senaste inspektion/underhåll (km), inspektion/underhåll krävs nu (booleskt värde).
För tåg (-noder):	Avgångstid (datum och tid), avgångsort, ankomsttid (datum och tid), distans (km), försening (min), inställt (booleskt värde), har avgått (booleskt värde)
För alla kopplingar (kanter till tåg)	Giltig (booleskt värde), aktuell (booleskt värde).
För verkstad:	Ort, öppettid, medeltid för verkstadstid (minut)
För depå:	Ort, öppettid.

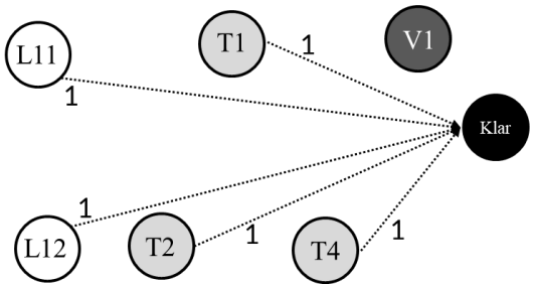
Handlingen som agenten utför är att sätta vikter på de kanter som diskuterades i stycket ovan om lokplanering modellerat som en graf. Givet dessa vikter genereras sedan en komplett lokplanering med så liten kostnad som möjligt. I ”Figur 8: Lokplaner specificerade av kantvikterna.” finns två exempel på hur olika vikter ger olika lokomlopp.





Lok L11 drar tåg T1 och sedan lok T4 innan det kör till verkstad V1. Lok 12 drar tåg T3 innan det kör till verkstad V1. Inga tåg inställda.





Lok L11 står kvar. Lok L12 står kvar. Alla tåg inställda.

Figur 8: Lokplaner specificerade av kantvikterna.

Sist men inte minst behövs en *belöning*. Belöningen, r, räknas ut med följande formel för den lokplanering som agenten genererat:

$$\beta = 1000n_f + 300n_c + t_d$$
$$\gamma = \underline{b} + n_a$$
$$r = \begin{cases} -\beta, & \beta > 0 \\ \gamma, & \beta = 0 \end{cases}$$

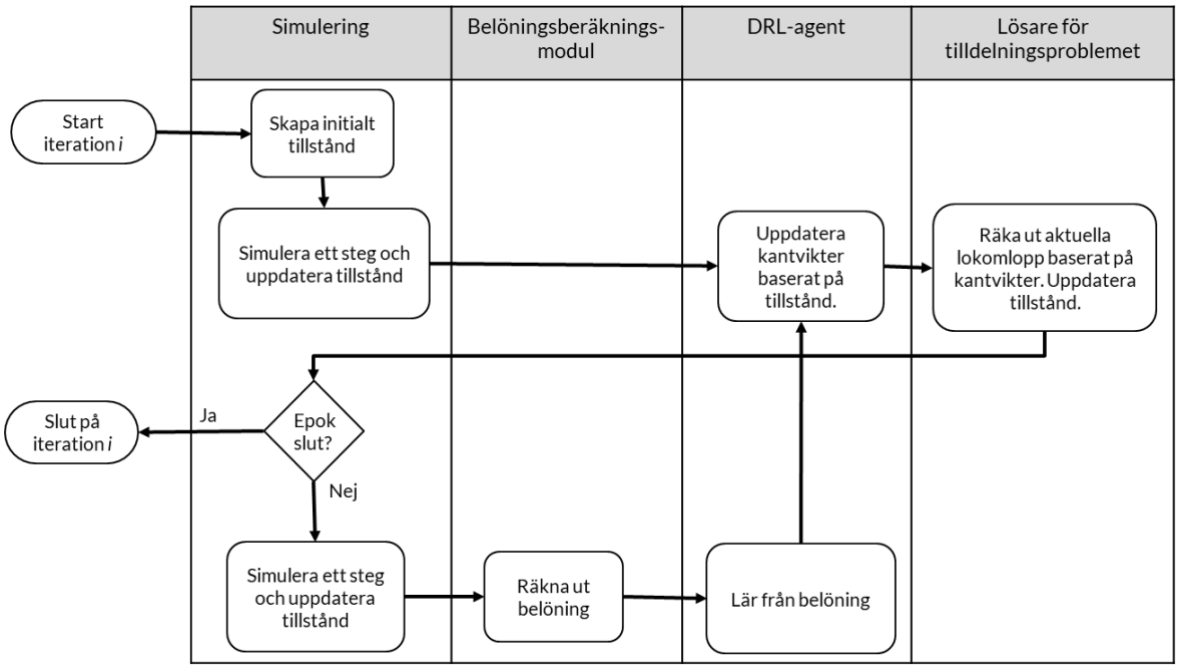
där n_f är antalet lok som fått körförbud, n_c är antalet inställda tåg och t_d är summan av förseningsminuter, $\underline{b}$ är den minsta buffertiden i någon koppling och n_a är antalet kopplingar i lösningen som är aktuella.

Med den här belöningsfunktionen straffas körförbud hårdast, följt av inställda tåg. Förseningar straffas proportionerligt med antal förseningsminuter. Förseningsminuterna ges av simuleringen (se nästa kapitel). Om lokplaneringen inte innebär något körförbud, inställelse eller försening så blir belöningen den minsta buffertid som finns i kopplingarna och antalet aktuella kopplingar som återanvänds. Buffertiden inkluderas för att få en robust plan, medan antalet aktuella kopplingar som återanvänds ska begränsa onödiga förändringar i planen.

4.3.2 DRL-träning genom simulering

För att träna DRL-agenten föreslås att trafiken simuleras så att följdförseningar kan uppskattas. I Figur 9 visar vi detta upplägg. Simuleringen kan vara allt från en väldigt enkel trafiksimulering där förseningar minskar i relation till tågens och kopplingarnas buffertider, till mer fullödiga simuleringar där både tågens fysiska egenskaper och tågklarerarens beslut simuleras. När det kommer till trafiksimuleringar finns t.ex. mikrosimuleringssystemen RailSys (https://www.rmcon-int.de/railsys-en/) och Open Rails (https://www.openrails.org/), eller makrosimuleringssystemet PROTON (Zinser,

o.a., 2019). För att bättre simulera tågklarerarens påverkan på förseningarna kan även optimeringsmodeller användas, till exempel forskningsprototypen M2. M2 använder optimering för att hitta en giltig tidtabell givet en målfunktion (Gestrelius, 2022). För att simulera en tågklarerare kan denna målfunktion till exempel vara att tåg ska ligga så nära sina ursprungliga tåglägen som möjligt, alternativt att minska förseningar. När man väljer vilken simulering som ska användas behöver man tänka på att (1) de aspekter som man vill att agenten ska beakta måste simuleras på ett tillräckligt bra sätt, samt (2) simuleringen måste gå tillräckligt snabbt. Simuleringen kommer köras i varje träningsiteration, och det krävs många iterationer för att träna agenten. Dessa två mål är ofta motstridiga, så det är en grannlaga uppgift att välja rätt simuleringsmetod och detaljeringsnivå. Om man väljer att använda en optimeringsmodell för att simulera tågklareraren kommer man behöva begränsa sökrymden på ett lämpligt vis för att hålla simuleringstiderna nere. Man kan till exempel begränsa hur mycket optimeringen får flytta på tåg som går i tid, eller kanske inte låta optimeringen flytta på dem alls. Ytterligare en metod för att simulera trafik är att använda någon av de AI-modeller som tränats för att uppskatta tågförseningar utifrån ett givet trafikläge, till exempel modellen från https://www.thetrainbrain.com/.



Figur 9: Hur de olika komponenterna samspelar.

4.3.3 Ett första försök

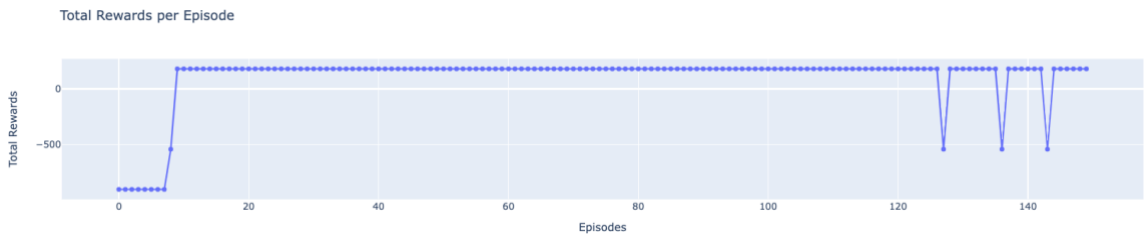
Som ett absolut första skott programmerades ett DRL-ramverk i Python och testades på ett mycket litet exempelfall bestående av två lok, tre tåg och en verkstad (samma som använts i Figurerna Figur 5: Lokplaneringsproblemet som en graf.-Figur 7: Matris för tilldelningsproblemet. ovan). Verkstaden antogs alltid vara öppen och inga kopplingar var satta som aktuella. Simuleringen var enklast möjliga och bestod endast i att propagera en initial försening via kopplingar och då också samtidigt minska följdförseningen med så pass mycket buffertid som fanns i kopplingen. Algoritm 1 är en pseudo-algoritm för hela ramverket.

Algoritm 1: Enkelt ramverk för en träningsiteration av DRL-agent

Input	:	Initialt tillstånd $\boldsymbol{I}$, DRL-agent $\boldsymbol{A}$**,** kopplingsgraf $\boldsymbol{G}$.
Output	:	Agent $\boldsymbol{A}$ som tränats i en iteration.
Steg 1	:	Skicka tillstånd $\boldsymbol{I}$ till DRL-agent $\boldsymbol{A}$ som uppdaterar kantvikter i kopplingsgraf $\boldsymbol{G}$.
Steg 2	:	Räkna ut omlopp baserat på kantvikter i kopplingsgraf $\boldsymbol{G}$, uppdatera tillstånd $\boldsymbol{I}$ med de nya omloppen.
Steg 3	:	Låt simuleringen uppdatera tillstånd $\boldsymbol{I}$ fram till nästa tågavgång och räkna ut straff och belöning, skicka till agenten $\boldsymbol{A}$.
Steg 4	:	Nästa tåg avgår. Uppdatera tillståndet $\boldsymbol{I}$ och kopplingsgraf $\boldsymbol{G}$ med tågets avgång.
Steg 5	:	**Om** alla tåg avgått: Returnera $\boldsymbol{A}$. **Annars**: Gå till *Steg 1*.

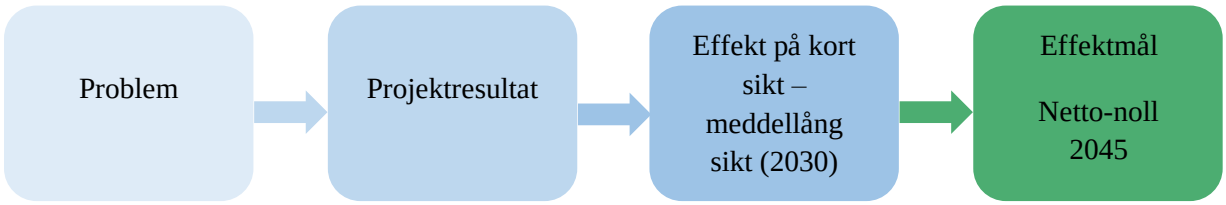
Notera att simuleringsstegen i detta ramverk ges av ståltågens avgångstider. Eftersom inga nya initiala förseningar uppstår efter den första så räcker så pass grova steg.

Flera olika träningsrundor gjordes för att försöka hitta lämpliga hyperparametrar och målfunktionsvikter. Modellen är inte stabil, utan lyckas bara ibland tränas till att ge en bra lösning. I träningsrundan som syns i Figur 10 ger agenten oftast lokomlopp som är optimala, dvs. där målfunktionen är positiv. Slutsatsen är att det går att programmera upp en agent utifrån det som beskrivits i rapporten. Det går dock inte att säga om ett sådant beslutsstöd kommer ge bra stöd eller inte. Den instabilitet som finns i vårt lilla exempel är inte acceptabel för ett riktigt verktyg, men den kan förklaras med att träningsdatamängden var allt för liten, och med att hyperparametrarna och målfunktionen inte kalibrerades tillräckligt. Även om instabiliteten inte är förvånande, så pekar den ändå på vikten av, och svårigheterna i, att kalibrera och träna en agent så att den blir användbar.



Figur 10: Den totala belöningen per iteration (kallad ”episode”) i en träningsrunda.

5 Nyttiggörande och nästa steg



Operativ lokstyrning är en komplex uppgift där DRL-agenter borde kunna användas som beslutsstöd. I SOL-förstudien har vi tagit fram ett förslag på hur en DRL-agent skulle kunna utformas och tränas för att utgöra ett sådant beslutsstöd, samt gjort ett första, mycket begränsat, försök till träning. Vi har visat att de steg som beskrivs i rapporten ger en agent som ändrar på lokomlopp. Det är dock fortfarande oklart om agenten kan bli ett adekvat beslutsstöd eller inte. För att svara på den frågan krävs en större studie. Projektets unikhet ligger i tillämpningen av DRL-agenter för lokstyrning.

Effektmålet för Triple-F är netto-noll 2045, och bättre beslutsstöd för operativ lokstyrning kan bidra till detta genom att öka resurseffektiviteten och förbättra leveranskvaliteten för godstransporter på järnväg, och således främja överflyttningen till detta energisnåla transportslag. Notera att en lyckad implementation av en DRL-agent för lokstyrning även öppnar upp för liknande agenter inom andra resursoptimeringsområden inom järnvägssektorn, såsom personalstyrning, vagnsbokningar och tågtrafikledning. Likaså har DRL-agenter stor potential att skalas upp och liknande agenter bör kunna tillämpas inte bara för Green Cargos resursstyrning utan även för andra godstågoperatörer nationellt och internationellt.

SOL är en förstudie, och för att nå de utpekade målen krävs ytterligare forskning samt slutligen, om lämpligt, implementering inom tågbolag. Ett första steg mot nyttiggörande är att sprida den kunskap som tagits fram i förstudien. Resultaten har presenterats för relevanta personer på Green Cargo, och under denna presentation diskuterades även vilken kompetens och förberedelse som behövs inom tågbolag för att de i framtiden ska vara redo att anamma AI-baserade beslutsstöd.

För RISE del har förstudien lett till att det nu finns kompetens och ett första ramverk för att träna en DRL-agent för lokstyrning. Detta stärker vår förmåga att bedriva ändamålsenlig forskning inom området, och underlättar vidare studier.

Förstudiens resultat kommer också presenteras för externa på Triple F-konferensen den 18e oktober 2023 samt, om möjlighet ges, på KAJTs höstkonferens. Den här rapporten bidrar också till kunskapsspridning.

Nästa steg efter förstudien är att fatta beslut om ett större genomförande- eller doktorandprojekt för vidare utveckling och implementering av den utvecklade metoden. Sådana diskussioner har påbörjats men inga beslut har tagits än.

6 Diskussion

Förstudien "Självlärande neurala nätverk för operativ lokstyrning" (SOL) undersökte möjligheten att ta fram ett AI-baserat beslutsstöd för resursstyrning för godstågsföretag. Specifikt har problemet med operativ lokstyrning studerats, men resultaten bör till viss del vara applicerbara även på vagnsstyrning och personalplanering.

Dataanalys visar att man på ett stort godstågföretag i Sverige, Green Cargo, jobbar mycket aktivt med lokstyrning, och loken bokas frekvent om. Det är ofta svårt att fullt ut förstå hur en ombokning av ett lok påverkar trafiken och transportkvaliteten, och därför skulle ett intelligent beslutsstöd vara av nytta. Ett av resultaten från SOL är en beskrivning av hur en AI-baserad agent kan utvecklas och tränas för att stötta en lokplanerare. Den föreslagna lösningen modellerar lokomloppsplaneringsproblemet som en graf och agentens handling är att sätta vikter på grafens kanter. Vi har också konstaterat att det finns data av tillräckligt god kvalitet hos Green Cargo för att kunna utveckla ett sådant system. För att uppskatta effekten av att implementera den beskrivna lösningen krävs dock att en större studie genomförs. Generellt kan dock sägas att bättre beslutsstöd för lokstyrning bör leda till bl.a.:

- Förbättrad resurseffektivitet och kostnadsminskningar inom godstransportssektorn i stort, och den inom järnväg i synnerhet, genom optimerad lokomloppsplanering och resursallokering.
- Minskade förseningar och bättre leveranskvalitet för järnvägstransporter genom snabbare och mer exakta beslut vid omplanering pga. störningar eller förändringar i trafikförhållandena.
- Ökad konkurrenskraft för järnvägen jämfört med vägtransporter genom minskade kostnader och bättre kontroll över resurserna.
- Minskade koldioxidutsläpp genom överflyttning av gods från väg till järnväg, vilket bidrar till att uppnå ett mer hållbart godstransportsystem.

För att ta nästa steg och förverkliga dessa effekter behöver följande åtgärder vidtas:

- Fortsatt arbete med att förbättra datatillgängligheten och säkerställa att nya lösningar är kompatibla med befintliga system inom järnvägssektorn.
- Vidareutveckling och träning av en effektiv PPO-agent för lokomloppsplanering och undersökning av hur den kan återanvändas och integreras för vagns- och personalplanering.
- Utmaningarna med att göra AI användbart och förståbart

Referenslista

Arulkumaran, K., Deisenroth, M. P., Brundage, M., & Bharath, A. A. (2017). Deep reinforcement learning: A brief survey. *IEEE Signal Processing Magazine, 34*(6), 26-38.

Flodén, J., Bärthel, F., & Sorkina, E. (2017). Transport buyers choice of transport service - A literture review of empirical results. *Research in Transportation Business & Management, 100*(23), 35-45.

Gestrelius, S. (2022). *Optimisation models for train timetabling and marshalling yard planning.* Linköping: LiU Tryck.

Green Cargo. (2014). *Det här är Green Cargo*. Hämtat från Green Cargo: https://www.greencargo.com/sv/om-oss/ den 10 5 2022

Lundström, S. (2006). *Godskunders värderingar av faktorer som har betydelse på transportmarknaden.* Stockholm: KTH.

Nelldal, B.-L., & Boysen, H. (2014). *Gränsöverskridande godstransporter på järnväg Oslo-Göteborg-Köpenhamn-Hamburg: Utmaningar och möjligheter.* Stockholm: KTH Royal Institute of Technology.

SciPy documentation. (den 31 07 2023). Hämtat från Optimization and root finding (scipy.optimize): https://docs.scipy.org/doc/scipy/reference/optimize.html

Zinser, M., Betz, T., Becker, M., Geilke, M., Terschlüsen, C., Kaluza, A., . . . Warg, J. (2019). PRISM: A Macroscopic Monte Carlo Railway Simulation. *Proceedings of The 12th World Congress on Railway Research (WCRR).* Tokyo.