

Version: 1.0
Datum: 2016-01-22



evry.com

Uppdragsbeskrivning

Predictions

EVERY Integration AB

Versionshistorik

Ändring nr.	Ändring datum	Förändringar	Reviderad av
1.0	16-01-22	Dokumentet skapat	Torbjörn Stake

Innehållsförteckning

1	Inledning	4
2	Bakgrund	4
<hr/>		
2.1	Prediktiv Analys, VACS och det svenska skogsbruket	4
2.1.1	VACS	4
2.1.2	Prediktiv analys	4
2.1.3	Det svenska skogsbruket	5
2.2	Problemformulering	5
3	Primära roller och kravställare	6
<hr/>		
3.1	Produktchef	6
4	Övergripande målbild	6
<hr/>		
4.1	Primära mål	6
4.2	Sekundära mål	6
5	Projektstyrning	7
6	Stöd från EVRY	7
<hr/>		

1 Inledning

Detta dokument är den överordnade uppdragsbeskrivningen för projekt *Predictions*. Syftet med projektet är att utvärdera plattformen Azure Machine Learning ur ett övergripande perspektiv. Syfte är också att belysa skillnader och likheter i de modeller som finns implementerade i plattformen, en studie och användning av programspråket "R" samt implementation av en eller flera ML-tjänster avsedda att användas för prediktiv analys i svenskt skogsbruk. Om tiden räcker görs också en klientimplementation där tidigare beskrivna ML-tjänster används.

2 Bakgrund

2.1 Prediktiv Analys, VACS och det svenska skogsbruket

2.1.1 VACS

Basen i EVRY:s skogliga affär är systemet VACS. VACS är ett system som utvecklats och byggts ut under 25 års tid. Det har genomgått ett antal tekniskiften genom åren och skapades ursprungligen för att köras i AS/400. Runt år 2000 lyftes systemet till SQL Server och klient i Centura. Då introducerades också namnet VACS (VirkesAdministrativ Client/Server). År 2006 tekniklyftes VACS klient från Centura till .NET och där befinner sig systemet idag, rent tekniskt.

Förutom anpassningar har VACS fått ett antal delsystem eller moduler utvecklade under systemets livscykel. Dessa moduler är mer omfattningsrika och kapslar lite större arbetsflöden eller funktionsgrupper.

Som ett nav för skogligt data agerar VACS också både datakälla och analyscentrum för skogliga transaktioner. Genom komponenten VACS datalager möjliggörs uppföljning som klassiskt beslutsstöd med analyser och rapporter riktade mot *historiskt* data.

2.1.2 Prediktiv analys

Prediktiv analys verkar framåtsträvande till skillnad mot klassiska beslutsstöd som uppvisar rapporter enbart på historiskt data. Att kunna förstå samband mellan datamängder och genom dem försöka säga något om sannolikheten för olika händelser inträffar framåt i tiden är ett bedömt framtidsområde inom fältet beslutsstöd som helhet. Med dagens VACS kan

ingen prediktiv analys göras och för att säkerställa systemets och modulen *datalagers* relevans i framtiden behöver sannolikt denna möjlighet tillföras.

2.1.3 Det svenska skogsbruket

Det svenska skogsbruket är en viktig näring för Sverige som nation och utgör en betydande del av vår BNP. Skogsbruket är kraftigt konkurrensutsatt ur ett globalt perspektiv och att kunna göra smarta analyser och framåtriktad planering är mycket intressant för att bibehålla denna konkurrenskraft. Dessutom påverkas skogsbruket av omliggande faktorer som makroekonomiska variabler, klimat och andra områden som utgör öar av s.k *big data*, som lämpar sig väl för *data mining* och prediktiv analys.

2.2 Problemformulering

Att ta in och utvärdera nya plattformar och teknologiska möjligheter är en utmaning för vilket bolag som helst. Framför allt när det gäller större koncept så som Azure Machine Learning och möjligheterna med prediktiv analys.

Det finns en stor mängd material och empirisk analys som behöver göras för att besvara frågan om vilka styrkor och svagheter som finns i plattformen. Att få detta presenterat i såväl rapportform som konkret implementerade exempeltjänster sparar EVRY både tid och pengar och minskar instegskostnaden och inlärningskurvan när, eventuellt Azure Machine Learning ska tas i bruk.

Följande problem finns för EVRY idag, relativt detta område:

- Vi har för närvarande inte kapacitet att själva inhämta en tillräcklig informationsmängd runt Azure Machine Learning för att bygga ett beslutsunderlag om plattformens lämplighet för oss.
- Vi har inte egen kapacitet för närvarande att utvärdera plattformen rent teknologiskt, så som att konstruera exempeltjänster och lära oss utvecklingsverktygen och behöver guidelines runt detta för att hjälpa oss igång.
- Vi kan för lite om programspråket R för att se möjlig potential i det som bas för program runt data mining.
- Vi vet för lite om hur enkelt eller svårt det är att sätta upp ML-tjänster och använda dessa från klient.
- Vi vet för lite om likheter och skillnader mellan de modeller som används i Machine Learning.

3 Primära roller och kravställare

3.1 Produktchef

Produktchefen behöver en fördjupad förståelse för Azure Machine Learning och en prototyp att använda i demonstrationssyfte för att:

- Kunna fatta välgrundade beslut om hur komponenten ska användas i EVRYs systemutveckling
- Hur svenskt skogsbruk skulle kunna tillämpa prediktiv analys
- Förstå inlärningstrappan för utveckling i programspråket "R"
- Förstå hur svårt det är att exponera ML-modeller som tjänster
- Förstå hur svårt det är att konsumera ML-modeller exponerade som tjänster

4 Övergripande målbild

4.1 Primära mål

- Att upprätta en rapport eller executive summary som beskriver plattformen Azure Machine Learning övergripande och dess lämpliga användningsområden.
- Att en eller flera ML-tjänster finns konstruerade i fullt och dokumenterade på ett sätt att våra utvecklare och konsulter kan använda detta för att öka sin egen förståelse.
- Att material tagits fram som belyser hur programspråket R förhåller sig till Machine Learning och en liten kunskapsbank om användande av detta språk.
- Att upprätta en rapport som beskriver ML:s olika modeller och dess skillnader och likheter tillsammans med exempel som visar på användning av dessa.

4.2 Sekundära mål

- Att få en sammanställning om hur tidskrävande utveckling i plattformen är samt vilka delar som kan förväntas vara komplexa att lära sig eller som vi behöver beakta vid implementation.
- Att bygga en klientapplikation som använder en eller flera publicerade ML-tjänster.
- Att det finns tjänster byggda med fokus mot svenskt skogsbruk och som skulle kunna hjälpa till med simuleringsmodeller så som:
 - En teoretisk klimatförändrings påverkan på ekonomiskt utfall i skogsbruket.
 - Simuleringar runt demografi och CRM "Sannolikheten att göra ett rotpostinköp i mellansverige med minst 7% vinstmarginal utan att först ha gjort en förkalkyl i VACS".

5 Projektstyrning

EVRY integration AB tillhandahåller en projektplats i Visual Studio Online och ett repository för att lagra källkod. Vi förutsätter att studenterna har tillgång till egna PC:s för utveckling, och med installationer av t ex Visual Studio.

Projektet drivs enligt en övergripande projektplan som studenterna tar fram i samråd med EVRY:s produktägare, och delas in i ett antal sprintar. Innehåll i varje sprint bestäms mellan team (studenter) och EVRY:s produktägare. Features bryts ner, läggs ut i Scrum Board och följs upp på regelbunden basis. Efter varje sprint genomförs en demo av implementerade funktioner.

6 Stöd från EVRY

EVRY:s produktägare bistår med stöttning runt projektplattformen och förtydligande/assistans runt kravformulering och precisering av uppdragsbeskrivning. Konsulter med expertområde skogsbruk assisterar studenterna med tänkbara användningsfall för modellering och analys samt dataextraktion ur EVRY:s egna system. Runt implementationsnära detaljer och konstruktion av klientdelar stöttar EVRY med programmeringskompetens.