

EXAMENSARBETE

2022-10-24

MASKININLÄRNING – ETT DATAKVALITETS-FRAMEWORK

Uddeholm är världens ledande tillverkare av verktygsstål för industriverktyg. Det är ingen tillfällighet. I varje steg av processen, från råmaterial till slutprodukt, är tekniken, kunskapen och produkt-utvecklingen i centrum. På digitaliseringsavdelningen söker vi nu studenter till examensarbete för att delta i detta arbete.

Bakgrund

Digitaliseringsavdelning på Uddeholm driver arbetet framåt för att utveckla hela Uddeholms organisation, från forskning och produktutveckling, till logistik och försäljningsprognoser. Avdelningen har bland annat expertis inom programmering, dataflöden, data science, maskininläring, big data och visualisering. Genom ett nära arbete med våra IT- och automationsingenjörer utgör vi kärnan i företagets satsning mot Industri 4.0.

Idag bygger Uddeholm data-pipelines med Apache Kafka, en komponent inom Apache ekosystemet för strömmande data i realtid. I takt med implementeringen av maskininlärningsmodeller som ett centralt element inom organisationen ökar också kraven på hög datakvalité för att säkerställa korrekta beslut från våra modeller. Syftet med detta examensarbete är därmed att definiera och implementera ett maskininlärnings-framework med fokus på datakvalité i strömmande data-pipelines.

**Uddeholms AB**

SE-683 85 Hagfors

Phone: +46 563 170 00 Fax: +46 563 174 00, www.uddeholm.com

Registration nr: 556046-2755 Registered at the Commercial Court: Hagfors, Sweden

Uppgift

Under vägledning av vår data scientist kommer examensarbetet att komplettera arbetet Uddeholm påbörjat tillsammans med Datavetenskap vid Karlstad Universitet inom ramverket för forskningsprojektet AIDA. De huvudsakliga uppgifterna kan brytas ned i två huvudsakliga komponenter: Definition och implementation.

Arbetet inom definition refererar till att skapa en klar begreppsbenämning på vilka datakvalitetsproblem som negativt påverkar maskininlärningsmodeller. Fenomen som konceptdrift inom deployment av maskininlärningsmodeller är något som Uddeholm har tidigare erfarenhet av, men begrepp som "data-smelling" och avancerad dataprofilering är nya termer som examensarbetet kommer utforska och definiera. Som en del av detta arbete kommer också alternativa strategier för maskininlärningsmodell-deployment behöva definieras för att hantera situationer då datakvalitetsproblem uppstår.

Den senare delen av examensarbeten, implementation, kommer bestå av att bygga ett framework för att upptäcka de definierade datakvalitetsproblemen samt att hantera maskininlärningsmodeller i kontexten av dessa datakvalitetsproblem. Arbetet kommer därmed säkerställa att prediktioner från Uddeholms maskininlärningsmodeller är tillförlitliga genom att använda "fallback"-modeller, korrigera datakvalitetsproblemen i realtid, etc.

Resultatet av examensarbetet förväntas leda till en genomgående analys av datakvalitetsproblem som negativt påverkar maskininlärningsmodeller. Arbetet förväntas också beskriva hur maskininlärnings-system bör hantera deployment av maskininlärningsmodeller för att minimera den negativa påverkan datakvalitetsproblem har på modellernas prediktioner. Slutligen förväntas arbetet också leda till en implementering av ett framework som hanterar just identifiering av datakvalitetsproblem samt deployment av maskininlärningsmodeller.

Genomförande

Arbetet kommer involvera både teoretiskt och praktiskt arbete, där studenterna förväntas undersöka officiell dokumentation, vetenskapliga artiklar, samt andra relevanta källor för att kunna definiera relevanta fenomen relaterat till datakvalitetsproblem. Studenterna kommer likväl behöva applicera och implementera dessa kunskaper inom ramverket av Uddeholm infrastruktur för att påvisa hur den nuvarande infrastrukturen skulle kunna förbättras med hantering av datakvalitet i strömmande pipelines genom ett generellt framework.

Avrapportering

Fullbordad rapport lämnas till handledaren samt presenteras muntligt under en sittning vid Uddeholm.

Avrapportering/avstämning under arbetets gång sker dygnsvis/veckovis till ansvarig handledare.

Sekretess

En sekretessförbindelse upprättas och signeras mellan studenten och Uddeholm.

Ekonomisk ersättning

Skattepliktig ersättning på 1 000 SEK/högskolepoäng utgår vid uppvisande av godkänt kursbevis, inlämnad och fullbordad rapport, samt genomförd muntlig presentation av rapporten.

Tidsram och ansökan

Examensarbetet är på 30 Hp med start i januari 2023. Ansökan görs via e-mail till nedanstående handledare.

Handledare/övriga kontaktpersoner vid Uddeholm

Namn (handledare): Erik Hallin
Telefon: 073 – 022 95 71
e-post: erik.hallin@uddeholm.com

Namn: Reine Lundin
Telefon: 072 – 544 34 09
e-post: reine.lundin@uddeholm.com

Uddeholms AB

SE-683 85 Hagfors

Phone: +46 563 170 00 Fax: +46 563 174 00, www.uddeholm.com

Registration nr: 556046-2755 Registered at the Commercial Court: Hagfors, Sweden