

Projekt AI & Kompany - Deep learning AI21

Syftet med det här projektet är att i grupp undersöka ett specifikt problem och hur det skulle kunna lösas med hjälp av deep learning. Detta innebär inte att ni behöver lösa problemet, utan gör efterforskningar, litteraturstudier, EDA, med mera för att planera, avgränsa problemet.

Notera att detta projekt är explorativt och möjlighet finns att ta och utveckla projektet vidare i kursen "Praktisk maskininläring". Det finns därmed inga förväntningar på att ni ska skapa färdiga modeller och "lösa" projekten inom tiden för denna kurs.

Slutprodukten kommer vara en A3-poster där ni i team presenterar under seminariet "vetenskaplig konferens" vad ni kommit fram till.

Case

Ni gör praktik på företaget AI & Kompany, ett konsultbolag inom IT som har för avsikt att verka inom data science och har kunder inom många olika områden. Företaget har en del pengar som de vill investera i ett inhouseprojekt. Tanken är att på sikt få fram ett proof-of-concept som de därefter kan gå och sälja in hos befintliga och nya kunder.

AI & Kompany har några projekt listade nedan som de tänker potentiellt kan bli business cases. De behöver hjälp från data scientist experter (alltså ni) att undersöka projekten.

Projekt: Forecasting temperatur

Gå in på denna sidan [temperaturdata SMHI](#), skriv in Göteborg och ladda ner all kvalitetskontrollerad historisk data. Undersök datasetet och hur man skulle kunna använda deep learning för att forecasta temperatur in i framtiden. Fundera på eventuella applikationer och business cases kring detta.

Projekt: Rekrytering

Kunden är ett rekryteringsbolag som sitter på många CVn. Just nu sitter de manuellt med många anställda som ska gå igenom CVn matcha till rätt jobb och skicka vidare. Ni får undersöka om ni mha dataanalys och maskininläring/deep learning kan hjälpa dem på något sätt. Datasetet hittar ni här på [Kaggle](#). Ett tips är att undersöka [Affinda API](#) testversionen som en start för parsing av CV. Fundera på hur en lösning skulle kunna bidra till att hjälpa ett rekryteringsbolag.

Projekt: Tag prediction

Kunder som driver bloggar och forum skulle vilja ha hjälp med automatisk taggning av olika topics. Undersök hur man skulle kunna gå tillväga mha dataanalys, maskininläring/deep learning för att automatisk taggning. Datasetet ni ska jobba med hittar ni här: [Stacksample Kaggle](#).

Projekt: Twitter sentiment

Det är många kunder som driver webshoppar som får en del reviews på olika produkter. Finns det något sätt att klassificera om en review är positiv, negativ eller neutral? Vad kan fördelen vara att känna till detta? Datasetet ni ska jobba med hittar ni här: [Twitter sentiment](#).

A3 poster

För att skapa A3-postern, kan ni välja vilket program ni föredrar, ta gärna inspiration från andra vetenskapliga posters. Gör ni i powerpoint ska ni gå in på Design -> Bildstorlek -> Utskriftsformat -> A3. Följande delar ska vara med, ni får ta med fler delar om ni vill.

- **Problem description**
 - avgränsa problemet - fundera på att det ska vara möjligt att inom ramen för kursen "Praktisk maskininlärning" ta fram ett proof-of-concept för en grupp på 3-4 personer.
 - **Literature review**
 - sammanfatta och extrahera viktigaste poängen från artiklar med liknande problem
 - **Data and EDA**
 - beskriv hur datan ser ut (visualisera gärna)
 - **Methods and models**
 - beskriv vilka metoder och modeller ni tänker använda
 - beskriv hur modellerna fungerar
 - **References**
 - referera till samtliga källor, ni kan välja mellan [APA](#) eller [IEEE](#), men var konsekvent.
-

Presentation

Ni har 15-20 minuter att presentera vad ni kommit fram till och ca 10 minuter för diskussion. Ni får använda valfritt presentationsverktyg som hjälpmedel. Samtliga ska presentera. Tänk på den röda tråden i presentationen.

Bedömning

Godkänt

- Uppfyllt kraven för A3-poster
- Utforskat datan och föreslagit eventuella lösningar
- Förklarar modeller, datan, problemet
- Samtliga presenterar