

Kursnamn: AI
Klass: PAI3
Termin: VT23

Telcco AB

INLEDNING

Bakgrundsbeskrivning

Telcco AB är ett fiktivt företag som har en utmaning med kundbortfall. Man vill identifiera de kunder som riskerar att falla bort. Dessa kunder kan genom riktade erbjudanden lockas att stanna kvar.

Själva datan är baserad på en publik Kaggle-tävling, men något modifierad. Den beskrivs här:

<https://www.kaggle.com/datasets/blastchar/telco-customer-churn>

ANVÄND DOCK BARA MIN DATA.

Varför ska ni utföra detta arbete?

Syftet är att visa hur man kan använda AI för att lösa ett verklighetstroget problem, genom att undersöka vilken modell som är bäst lämpad för denna uppgift. Visa att du kan göra den analys (EDA) som behövs innan man använder ML modeller, och att du kan undersöka olika variationer av parametrar för att utveckla och optimera respektive modell.

Vad ska ni leverera?

Uppgiften lämnas in som en Jupyter notebook (en .ipynb fil), där stegen beskrivs och där all kod, resultat, text (här ska du använda dig av markup för att skapa titlar och dyl.) och grafer finns samlade. Varje del ska vara tydligt markerad med en titel, med betoning på presentation.

Filen ska heta **FÖRNAMN.EFTERNAMN.ipynb**.

All kod och dokumentation bör finnas i notebooken.

ER PROJEKTUPPGIFT

Vad ska ni göra?

Du ska hitta en lösning som bygger på observationer som intäkter från kunden, antal produkter, och andra data om kunden. Man vill hitta kunder som riskerar att falla bort.

Hur löser ni uppgiften?

Uppgiften är baserad på publika dataset om kundbortfall, s.k. *churn*.

Du ska presentera en lösning med verktyg du vill använda.

Du ska utforska, analysera och presentera datan.

Förbereda data; t.ex. fixa inkorrekta eller ofullständiga värden, omvandla kategorisk data och gör de transformationer som behövs för användning i en modell.

Utforska olika modeller och utveckla dem. Visa resultatet tydligt.

Följande termer är en guide till vad som ska vara med.

- Träning och prediktion.

- Mätvärden.
- Model selection and accuracy.
- Bias-variance trade-off.
- Model hyperparameter optimization.
- Supervised/ Unsupervised.
- Neural Networks.

Varje graf ska ha en

- titel
- notering vad x- och y-axeln visar
- annan dokumentation som är nödvändig för att förstå grafen.

Ha med länkar av de källor du använt dig av för att lösa uppgiften.

Struktur för arbetet

Uppgiften är individuell.

INLÄMNING OCH REDOVISNING

Inlämning

Projektet skall vara klart och färdigställt för bedömning senast den **14/4** kl 23:59. Skall lämnas som en .ipynb fil genom ladda upp i learnpoint.

Redovisning

BEDÖMNING OCH ÅTERKOPPLING

Bedömning sker med följande betygsriterier

För godkänt (G) på projektarbetet skall följande krav uppfyllas:

Adekvat undersökning av data med kommentarer.
Förklaring av vad för typ av AI är bäst lämpad för denna data.
Undersökning och jämförelse av **minst två** fungerande modeller.
Variera olika parametrar och diskutera deras effekt på noggrannheten.
(Experimentering.)
Diskutera även själva noggrannheten och beskriv valet av utvärdering av modellerna.

Presentationen är tydligt uppdelad i sektioner med relevanta titlar.
Grafer har en titel och x- och y-axel kommentarer så att det är lätt att förstå vad den föreställer.

För väl godkänt (VG) på projektarbetet skall dessutom följande krav uppfyllas:

Undersökning, experimentering och jämförelse av minst fyra modeller.
Presentera bästa möjliga modell och argumentera varför den är bäst.
Förklara även vad resultaten innebär så att en icke teknisk åhörare förstår.

Återkoppling

28/4.