

# Prediktion av bokföringskonto med en multinomial logistisk regressionsmodell

Henri Jäderberg\*

Juni 2017

## Sammanfattning

I detta kandidatarbete konstrueras en modell som predikterar vilket bokföringskonto en kostnad bokförs på. Den modell som används är en *Baseline-Category Model*, som är en generalisering av logistisk regression från två till flera klasser. Vid skattning av parametrarna används regularisering av typen Ridge regression. Modellsektion görs genom korsvalidering där förlustfunktioner jämförs. Den slutgiltiga modellen utvärderas med avseende på hur väl den presterar då den anpassas på helt ny data. Modellen predikterar 57.02% av observationerna rätt. Detta kan jämföras med att alltid gissa på det mest förekommande kontot, vilket uppskattningsvis predikterar 18.41% rätt. Prediktionsförmågan varierar dock beroende på vilket konto som observationen faktiskt antar. Mer data har troligtvis endast potential att öka andelen observationer som modellen predikterar korrekt till en nivå på drygt 65%. För att nå högre behövs nya prediktorer.

---

\*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige.  
E-post: [h.jaderberg@gmail.com](mailto:h.jaderberg@gmail.com). Handledare: Tom Britton och Benjamin Allévius.