**Project 3 Released**

#601

Project 3

**Project Description**: [[HTML](https://liangfgithub.github.io/Proj/F23_Proj3.html)] [[PDF](https://liangfgithub.github.io/Proj/F23_Proj3.pdf)] **Due**: Monday, December 4 **Dataset**: Available on Coursera/Canvas

Library Restrictions: TBA

**Rubric**: 7.5pt in total

* **2.5pt** for report
* **3.5pt** for performance (measured by the minimal value of AUC over the five test datasets)
  + 1pt: worse than 0.90
  + 2pt: [0.90, 0.94)
  + 3pt: [0.94, 0.96)
  + 3.5pt: 0.96 or better
* **1.5pt** for vocabulary size: A model based on a smaller vocabulary size is easier to interpret.
  + 0.5pt: vocabulary size in (2,000 3,000]
  + 1pt: vocabulary size in (1,000, 2,000]
  + 1.5pt: vocabulary size <= 1,000
  + Each entry in your vocab could be a word (including numbers) or a phrase, such as "worst\_movie".