

Leadbook ML Challenge

Problem Description: In Leadbook we have millions of contacts and thousands of new contacts added daily. One of the important tasks is to classify the departments/categories from Job title of contact. In this challenge the job titles should be classified into respective departments/categories using NLP/ML techniques.

To Do: Please develop a model which can classify the given job title into departments/categories. A job title may have more than one department/category or none.

- 1) Do pre-processing if required.
- 2) You can choose any techniques ,algorithms or existing models for classification.
- 3) You can use cascading or ensemble methods if required

Given: jobtitles_all.csv - 84k Job titles departments.json - Each department has sample keywords, which can be used for learning,clustering,prediction etc..

Note: Ignore the job titles 1) which only have seniority like CEO,Senior manager..etc 2) Junks like zzz,sab 3) Non English

Evaluation: We are looking for the below metrics.

- 1) How did you approach the problem
- 2) Why did you choose this method/algorithms. Please describe about your epochs,activation function and other metrics if used
- 3) What is your accuracy, how did you calculate it and what are the future works needs to be done to improve it

Please submit your answers,codes in **git repo**. Explore and Enjoy the task. We would like you to come up with **estimated time to complete**. Let us know if you need further information.