**Recommendation System For Domain and Events**

Made by : Elita Menezes

(Machine Learning Intern)

Project Manager : Jayanth G S

### Introduction

Due to lack of awareness one often misses events of interest related to their domain. Cloud Counselage also receives such invites for events in multiple domains which need to be forwarded to people with relevant interests.

The existing system includes sending the invites of each domain and event to every individual to ensure no one is left out. However, if a person receives emails about events he/she is not interested in over a period of time they tend to ignore the mails of the sender, thus missing the opportunity.

### Project Overview

With drawbacks of the current existing system, the project aims at building an automated recommender which outputs a csv of names to whom the event must be recommended.

Once a new event is entered, the system will then extract the information like the domain name and the event from the text. Then it will go through the csv to find all names that are interested in the following and produce a csv of the same.

The concept used for extracting information is Named Entity Recognition.

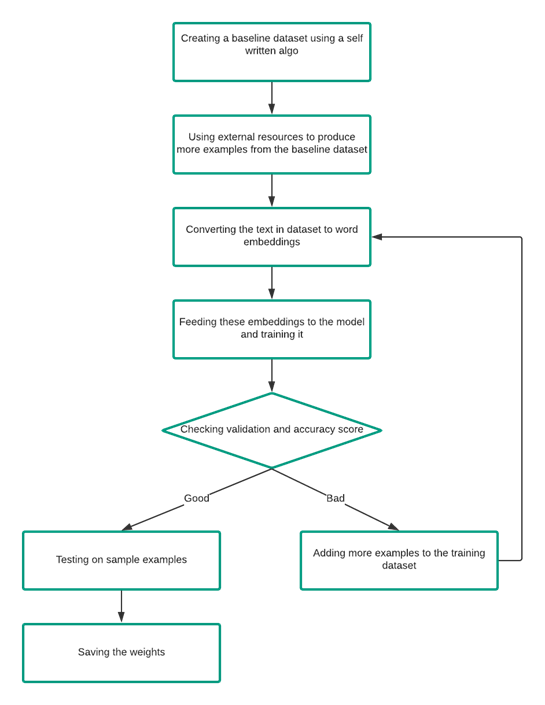
It classifies named entities in the text into predefined categories such as name, person, organization etc., and in our problem statement into domain and events.

### Implementation details

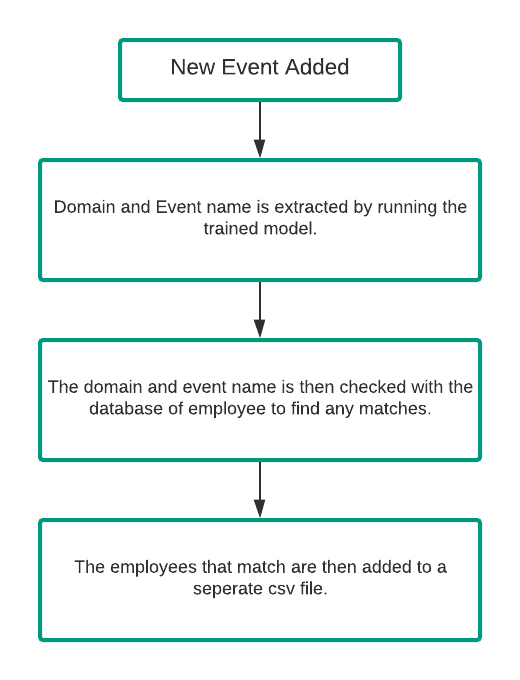
* 1. Design Model

The project is divided into 2 main parts:

1. To train the model to classify the events and domains.



2. Using this trained model to classify the same, get employees matching the criteria from the database and produce a csv file of the same.



3.2 Final Solution:

First I created a dataset of 75 sentences by taking inputs as different post from LinkedIn. These sentences were self-annotated using a python script that I created. I then trained a Named-Entity-Recognition model using these sentences for 500 iterations and tested it on a smaller dataset. It extracted about 60-70% of the words correctly, therefore as a failsafe I made use of a python library difflib which helps in providing closest matches to a word in a sentence.

Once the domain and event names were extracted these were then checked across the given employee dataset and it gives a dataframe of the input with comma separated values of the employees.

3.3 Improvements:

Currently this solution takes an input as a text or string; however people promote their events as an image. Therefore, a better approach would be to extract the text from image using OCR techniques and then using this text to further extract the domain and events.

### Results and Conclusion

Recommender engines are being widely used in a number of websites like YouTube, Google, Facebook, and Amazon. As part of Cloud Counselage considering active hiring of employees it will be good to provide these interns with an opportunity to grow by constant updates of the latest events of their field of interest.

The recommender system that I created using spaCy, helps in the task of classifying these inputs. The model has been trained on approximately 70 sentences taken from a popular website LinkedIn to provide inputs similar to the actual data. It shows a loss of 9.34 after 500 iterations of training. As part of a fail safe, it uses difflib library to check for words with similar meanings.

The recommender engine is at its initial stage and needs a lot more training data for it to be successfully implemented.

### Reference

* [BERT Explained: State of the art language model for NLP](https://towardsdatascience.com/bert-explained-state-of-the-art-language-model-for-nlp-f8b21a9b6270)
* [google-research/bert: TensorFlow code and pre-trained models for BERT](https://github.com/google-research/bert)
* [What is BERT | BERT For Text Classification](https://www.analyticsvidhya.com/blog/2019/09/demystifying-bert-groundbreaking-nlp-framework/)
* [EntityRecognizer](https://spacy.io/api/entityrecognizer)
* [Custom Named Entity Recognition Using spaCy | by Kaustumbh Jaiswal](https://towardsdatascience.com/custom-named-entity-recognition-using-spacy-7140ebbb3718)