Cloud Counselage Pvt. Ltd.

Machine Learning Internship Live Project Report

(Employee Recommendation System Using Machine Learning)

Intern Name: Ramakant Shrihari Vaidya

Project Manager: Jayant G. S.

Project Timespan: 7 Weeks

Acknowledgment

This project is made by idea given from Cloud Counselage Pvt. Ltd. developed as an intern for machine learning for 3 months. Project recommends employees based on their preference or area of interests. I am thankful to Cloud Counselage for giving me this opportunity to have a project based on their idea and it was a great pleasure to work as an intern with company. Thanks to company for guiding in the project and helping every time.

Abstract

This project focuses on a recommendation of employees based on their interest. This system draws data from given input and checks whether domain and event in input matches with any of the employee or not. This project may help both company and employees for selecting employees and getting job or work in field in which employee is actually interested. Project uses bag of words approach with Logistic Regression and forms a CSV file as an output file for given input file.

List Of Abbrevations

ML : Machine Learning

NLP : Natural Language Processing

Problem Statement:-

Create a deliverable which will provide employee names based on the interested domains and events for the input event. If event and domain matches with employee then system should return name of employee.

Employee Recommendation system

Using Machine Learning approach

As name suggests this project focuses on recommending names of employees who might be interested in the work available. This can help to both company as well as employee as this gives company the names of interested candidates and employees gets chance to work in their desired domains. If multiple employees are interested in the same domain with interested event then system provides names of every employee.

Project can help a company to filter employees and not hiring not interested employees in any domain or giving interested domain if possible. The machine learning approach makes it easier and gets more perfect output and increases accuracy of project.

Project uses Logistic regression and bag of words for the recommendation purpose and then file handling for storing output to CSV file format. Libraries like pandas and numpy are used for some basic operations.

* Bag of Words:

Most of the ML algorithms works on numerical data and not on text data. So converting data into numerical form helps lot to train a model or get prediction from any model.

Bag of words uses similar way it creates a list of words such that it gives all names of domains and events which helps lot to train any model. This helped lot in the project as all data provided was a text data and input was also a text data. So this approach converted text data into numeric data.

* Logistic Regression:

As converted data was in 0/1 format predicting event and domain. Then these domains and events are predicted using Logistic Regression system can predict domain and event. These predictions can be used for predicting employee name as employee needed to selected having interest in domain as well as event.

* File Handling:

File handling is used to store output in csv file. The output is stored in dataframe and by reading dataframe data stored in CSV file. writerow() function is used for writing data in file row by row till data ends in dataframe.

* FlowChart:

Input available?

Return List with all employees applicable

Return Empty List

Domain = other

Event

domain

Apply Logistic Regression

Bag Of Words

Input

True False

False True

True

Conclusion:

This is clear from above report that bag of words and logistic regression is useful way to get output when dataset is in text format. Converting text to number and predicting from numerical data becomes much easier than directly dealing with text format data. This is not unique way but one of the effective way for this project to get desired output.