



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

TEAM DYNAMICS INTELLIGENCE SYSTEM

¹Manish Kumar Rohilla ²Prince Kumar ³Kaptan Yadav ⁴Tejal Batra ⁵Shubhita Menon

¹kotamanishrohilla@gmail.com ²princekumargaya890@gmail.com ³kaptanyadav007@gmail.com ⁴tejalbatra4@gmail.com

^{1,2,3,4}Student, ⁵Assistant Professor, School of Computer Science & Engineering

Lovely Professional University, Punjab, India

Abstract: The purpose of this research paper is to revisit the idea of online learning, freelancing and project out-sourcing. We have studied some literature to understand the basic concept behind, how communication through internet can help every individual to fulfil their demand of every kind and identify the problems in every dimension of the process. We are trying to propose a platform where you can team up with experts and work together to achieve your goals.

1. INTRODUCTION

Today in the era of technology and internet every internet user irrespective of their profession need help of technology. Some wants to have a career in technology, some wants to have their own start-up with their unique and innovative ideas. For every type of need we are trying to propose a solution and bridge the gap between the experts, help-lookers, businessmen and learners. The way of communication has changed with the evolution of technology. The purpose of this project is to make sure that Internet and Technology is used in the most efficient and optimized way. This project is also to provide opportunity of every kind to the users and handlers both.

2. SCOPE

We have professionals around us of every field but still we try hard to look for resources everywhere to learn and improve our skills. Some opts for courses, some goes to coaching and some opts for private labs. The problem with online courses and videos is, not every time we get what we are looking for. And the content sometimes gets outdated because of upcoming technologies, techniques or platforms. It's hard to keep up the pace of courses content update according to new demands and trend in market. In case of new startups, it is very hard for anyone to reach out to mass population, but a marketing guy knows the strategy to reach out and make the startup successful. That marketing guy looks for various ways of doing that which includes apps, advertisements and making web pages. But a marketing guy cannot make an app or a web page so for that we need a computer science engineer, for each kind of work we need a professional and someone with experience. In case of a student or learner, if someone wants to learn something or need help with some project but they are unable to get sufficient help, for them also this platform comes as a boon.

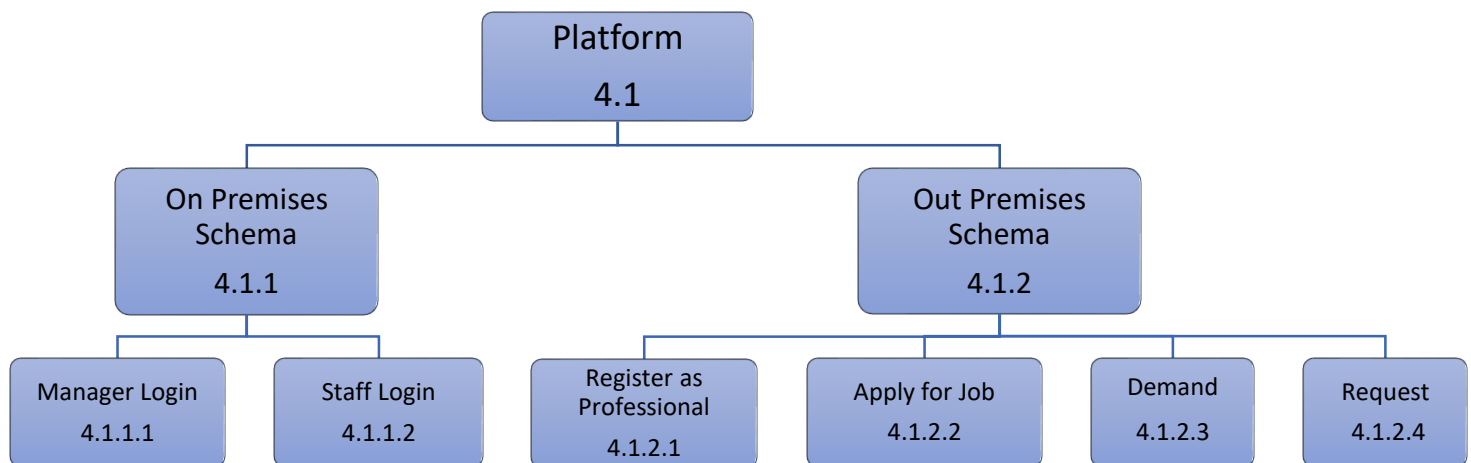
3. OBJECTIVES

We are trying to work on a simple all in one platform where the applicant just input their demand and we will provide direct bridge to reach out to professionals of their choice. This is not only for learners but also for any person who wants to work on some project or want to build something. For every kind of requirement there is one platform to reach out to. Professionals or freelancers looks for project to improve themselves and learners look for guidance.

4. MODEL EXPLANATION

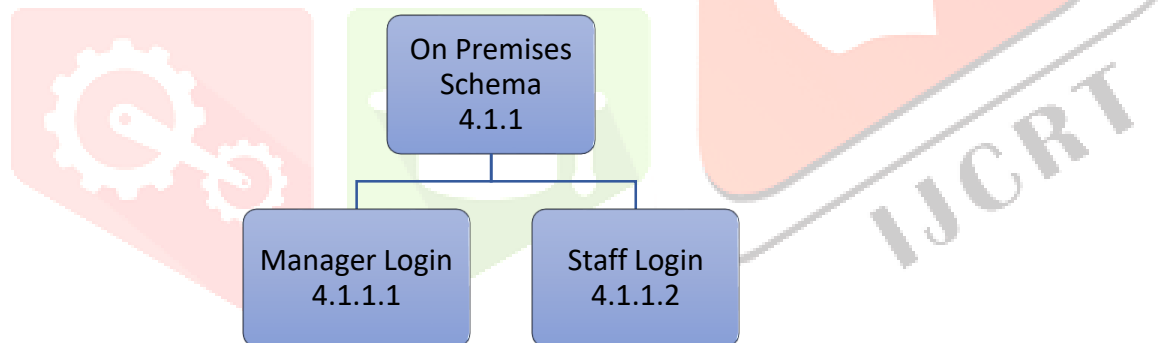
4.1 Basic Hierarchy of App

To make it easy to understand what the basic structure of this platform is and how it is connected to every module, how at every level of hierarchy the task is divided and how authorization varies between different kind of users. We are trying to explain it through the hierarchy chart. Chart has 3 levels. 1st is the platform itself, 2nd level divides the platforms between 2 parts which separates the user and service provider, 3rd level briefs about the services, tasks and responsibilities of the platform.



4.1.1 On Premises

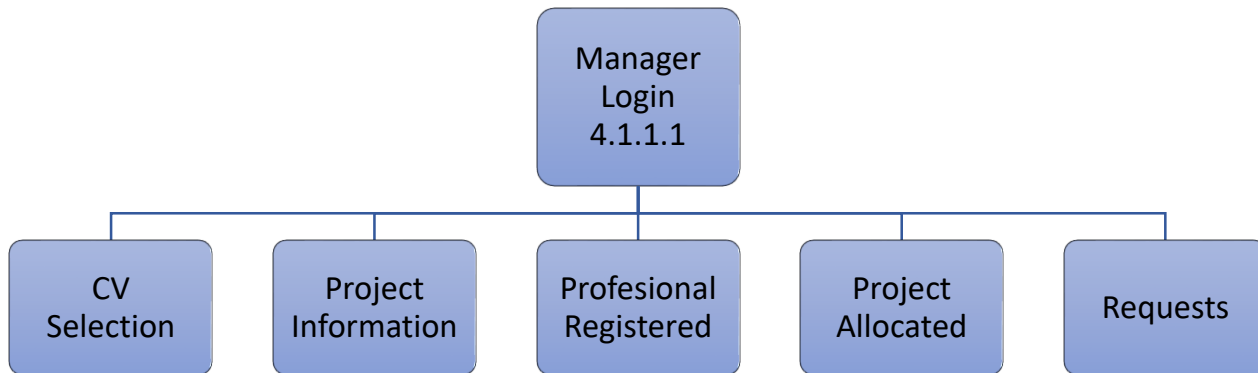
It is further divided into 2 parts, Manager login and Staff login. Authorization for manager login is more as compared to staff login. Staff Login do not have the authority to short list resumes and call for interview. Staff Login can only see the number of demand and the projects whereas admin login can look at the demand and the professionals selected under any project. Overall, On Premises schema revolves around the team and employees who are handling and maintaining the platform.



4.1.1.1 Manager Login

It will have CV selection module, Requests, Project Information, Professional Registered and Project Allocated.

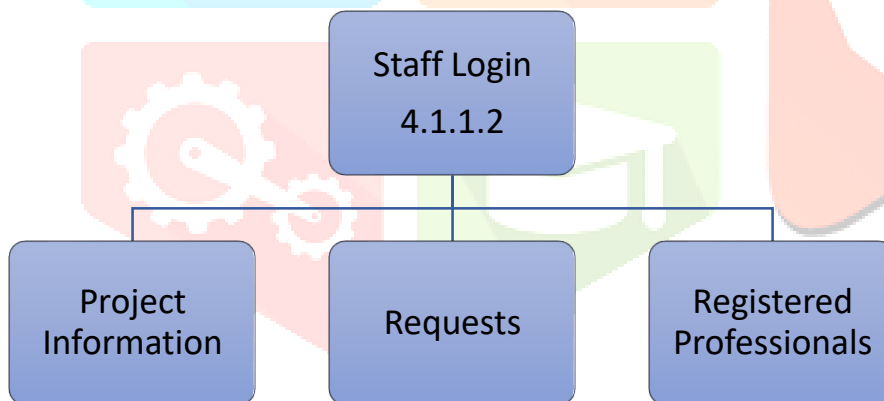
1. *CV Selection Module* will have lists of all CVs submitted through Job Application portal and the manager have the authority to pick any CV from the portal and evaluate if the applicant should be called for interview or not. For CV evaluation we have trained a Machine Learning Model using Naïve Bayes and Count Vectorizer.
2. *Project Information* module will have the list of projects or demands registered through Demand portal.
3. *Professional's Registered* module will list the Professional's registered from Register as Professional portal. Manager can also delete any professional from the Database if requested by the professional or if there is complaint towards that Professional.
4. *Project Allocated* module will list out the projects along with the Professionals selected in project.
5. *Requests* module will take care of the request submitted by professionals to deregister or to address complaints filed.



4.1.1.2 Staff Login

It will have Project Information, Requests or Complaint module, and Registered Professional's Information.

1. *Project Information in staff login* is as similar to Manager Login but staff members don't have the access to know the information about which professional is working under which project.
2. *Requests or Complaint* module will have the request submitted by customer or professional, they will attend those requests and pass on to manager after evaluation for action.
3. *Registered Professional's Information* will allow them to look at all the professionals registered.



4.1.2 Out premises

This module have 4 further divisions, which are Register as Professional, Apply for Job, Request and Demand.

4.1.2.1 Register as Professional

This module of out premises will allow any professional to become a member of the team by submitting some details and their resume. Their resume will also go under machine learning model for evaluation and on passing the evaluation they will be registered.

Every registered professional will be provided with Unique Id and this Unique Id will be required in case of any complaint or request made by either customer or professional respectively.

If in future, they want themselves to be deregistered then they can put the request for the same with their Id provided in complaint module.

4.1.2.2 Apply for Job

This module will allow anyone to work with the platform's handler team which will manage the communication and connection between the customer and professional to make it more user friendly, easy to use and make any user's professional life hassle free.

This module will take some inputs from the applicant along with their cv and only manager will evaluate it at the backend through trained Machine Learning model to check whether applicant should be called for interview or not and on the basis of applicant's input Machine Learning model will also predict whether the applicant is selected for staff post or manager post.

4.1.2.3 Demand

This module will take necessary input from the user to shortlist the professionals registered based on those input. Then they will get the list of shortlisted professionals and the user can select any professional from the list and submit their request.

Then the Manager will inform the user regarding the request submitted, staff do not have this authority to see the list of shortlisted professionals.

4.1.2.4 Request

This module will take inputs from the submitter. This module will also work as complaint page. This page will handle 2 tasks like request from the professional for deregistration and request or complaint from the user who have submitted request for professionals. In any of the case the unique id provided to them will be required for successful submission of request or complaint.

5. TECHNOLOGY USED

For making the platform interface we will use:

- Flask- Flask is third party python library used for developing web application, it is a web micro-framework dependent on jinja engine for template rendering and Werkzeug WSGI tools.
- Jinja2- It is a templating language that supports python. It manages markup formats from XML, HTML returned to user via HTTP request.
- HTML- A markup language to give structure platform.
- CSS- A web page styling tool to make webpages look attractive.
- Bootstrap- To make webapp responsive and adaptive.
- WTForms- To take user input for processing.

For Databases and file storage we will use:

- MongoDB- To store the information of every user like their IDs, password etc.
- Json (JavaScript Object Notation)- To store informational content or records.

For Machine Learning models that will help in shortlisting CVs and Professionals according to demand we will use Jupyter Notebook and Python Libraries listed below:

- Sklearn- Library for Machine Learning.
- Genism- Natural language processing, word embeddings and information retrieval.
- Count Vectorizer- To tokenize text and store it in form of sparse matrix.
- Regex- Regular Expression to extract certain type of patterns and replace.
- NaiveBayes MultiBinomial algorithm- Classification Machine Learning Model.
- Matplotlib- To plot graphs and figure to keep track of performance of our trained model.
- Word Cloud- To present frequently used words in any document in an attractive manner.
- Pickle- To save trained machine learning model and load it whenever needed.
- Seaborn- To plot graphs and figures.

To make our platform interactive and allow external software to access data we will use Postman API (Application Programming Interface). It is very useful to let user and system communicate and response to each other's request.

6. EXPERIMENTAL WORK

6.1 Algorithm used for Demand Matching.

For matching and allocating projects to the employees based on location and skill demand input, we will use Longest common subsequence algorithm and Regular Expressions (regex).

But we have a challenge to use it under our terms to match some particular words exactly.

For Example: For "Java" LCS can give perfect matching ratio 1 for the following:

Jacob Vallera ---> JAVA

**To overcome this problem, we will use some regex techniques to make string input converted in such a format that it will not create such problems. We will replace every non word character like space, commas, full stop, etc by using (\W) pattern and re.sub () function, then we will use “.split()” function with delimiter as “,” to separate it in form of lists.

We will perform these operations on both sides, on our employee’s dataset as well as user input.

After converting these both strings to list. We will execute Longest Common Subsequence on every element as reference for comparison w.r.t to other list.

We will put a counter to count where our elements of list matches with user inputs we will increment the counter by 1.

Then we will find percentage of similarity between inputs by using the following formula:

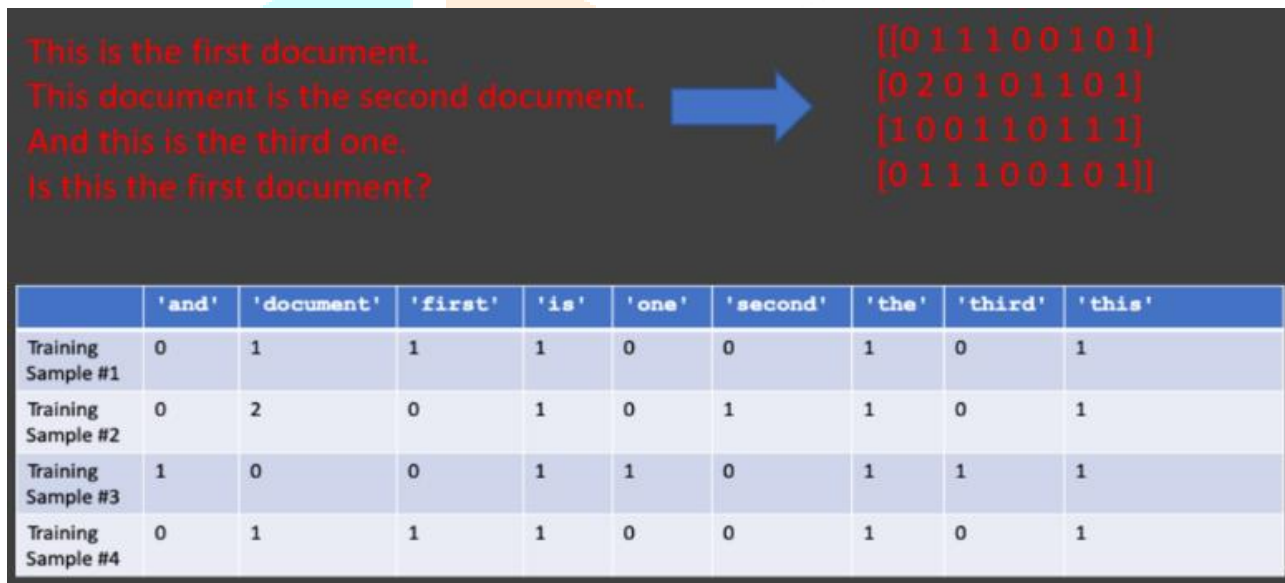
$$(2 * Counter / ((length of user_input_skill_list) + (length of professional_skill_list))) * 100$$

The outcome of the calculation will help in deciding that if the professional in check should be shortlisted or not.

6.2 For CVs Selection

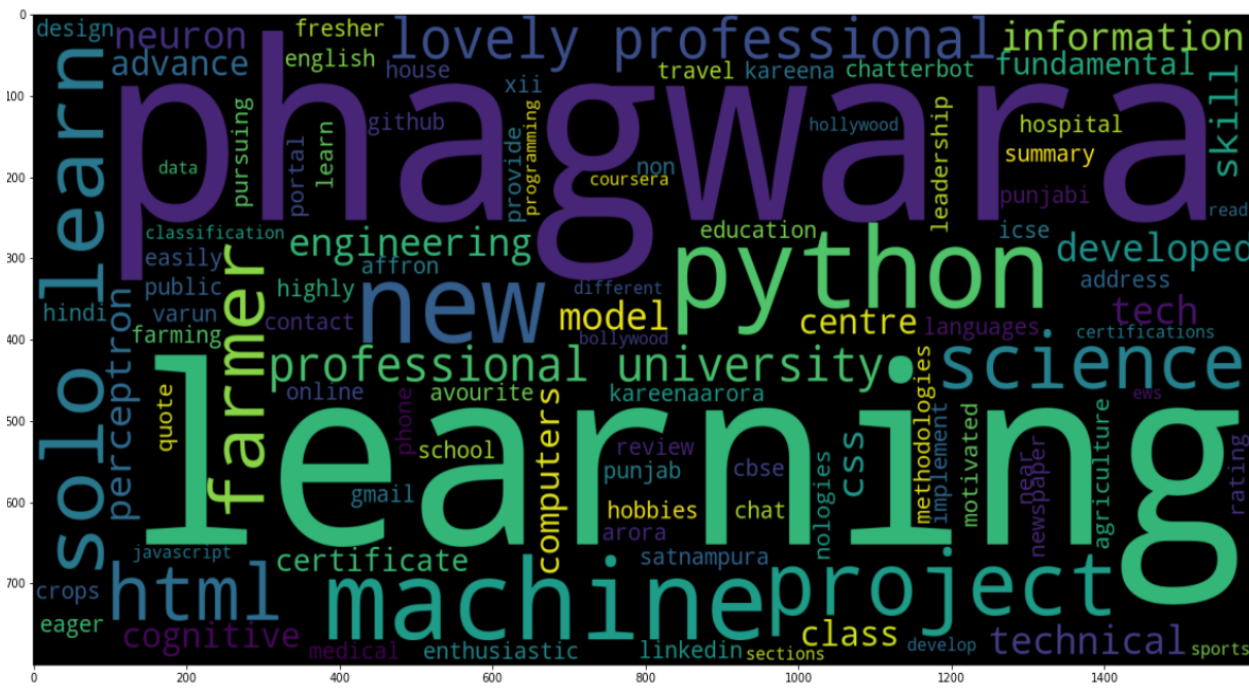
Every submitted CV will go under text extraction phase, every text then will be processed with the help of **Count Vectorizer** to tokenize them according to the vocabulary stored already from previously processed resumes. Then this tokenized text will be transferred to machine learning model trained on dataset from Kaggle.

6.2(a) Count Vectorizer Basic Working



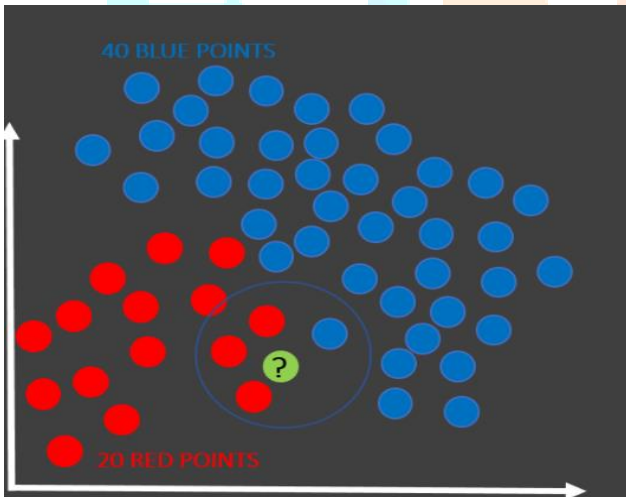
Before tokenizing the extracted text, we will do some pre-processing, which includes removing punctuations and unnecessary words (stop_words) from the text like “as, are, is, what, etc”.

We can also see the words most frequently used in the resume using word cloud it will look like as shown in plot 6.2(b)



The machine learning model we trained use **Naïve Bayes Algorithm**. This Algorithm is a multiclass classification algorithm, the class of every data point will be decided by the distance metric. In our case we will use binomial classification to decide whether to shortlist the resume or not we need only two instances, flagged or not_flagged.

6.2(b) Naïve Bayes basic Intuition



As it can be seen in the plot 6.2(b), we have data divided in 2 categories Red and Blue. And the Green Datapoint with Question mark(?) sign in it is not grouped in any class.

As the number of Blue points are more as compared to Red points the probability of it being grouped in Blue is more as compared to being grouped in Red.

- There are **40 Blue points** and **20 Red points** out of total **60 points**.

$$\text{Prior Probability for Red} = (\text{Number of Red points} / \text{Number of Total points}) = (20/60)$$

$$\text{Prior Probability for Blue} = (\text{Number of Blue Points} / \text{Number of Total points}) = (40/60)$$

Let's name Green unclassified point as X.

- We draw a circle around X and then we calculate the number of points in the circle belonging to each class label.
- For the data point X if there are more Blue point in the vicinity then it is more likely that X will be classified as Blue.
 $\text{Likelihood of X being Blue} = (\text{Number of Blue points in the vicinity} / \text{Number of Total Blue points}) = (1/20)$
- For the data point X if there are more Red point in the vicinity then it is more likely that X will be classified as Red.

Likelihood of X being Red = (Number of Red points in the vicinity/ Number of Total Red points) = (3/20)

- Then posterior probability is calculated:

Posterior Probability of X being Red = (Prior Probability for Red Likelihood of X being Red) = (20/60) *(1/20) = 0.016*

Posterior Probability of X being Blue = (Prior Probability for Blue Likelihood of X being Blue) = (40/60) *(3/20) = 0.09*

Result: The data point X will be grouped in **Blue class** because the Posterior Probability of X being Blue is greater than Posterior Probability of X being Red

7. CHALLENGES

There are several challenges we faced in implementation of this project. The main challenge was to make the model logical and easy to understand. Our platform is unique in many aspects and to give user the best experience we are integrating machine learning models. For training those models we lack datasets and without datasets it is nearly impossible to implement or train any machine learning model. We tuned the model and edited existing datasets in such a way that it decreases randomness, independency in our datasets and make it dependent or relational. We studied every possible technique to manipulate the working of model in such a way that it can be used to fulfil, what the platform demand. In the backend we used every optimized algorithm and library to make the platform user friendly and fast. We made sure that the authorization, duty and access schema difference is maintained between different kinds of every user.

8. CONCLUSION

This platform is for erasing the gap between every professional out there in the market and every person who are searching for resources to get their requirement fulfilled. The platform staff have very little part to play and most of the communication will be between users and professional members. By providing them one-in-all platform we are saving time, money and resources. No matter what the user wants to do, user don't need to think about where and how you will get help. You just visit the web-app input your demand select the professionals of your choice get their contacts and achieve your goals. If you are a professional then visit this webapp register as professional member, input required information, upload your resume and get ready to work, implement unique ideas and improve your skills. It will help in widening the learning spectra for everyone and will boost business and market. No one will spend their time on planning how and get tired, you just need to utilize your time and energy in implementation without any hassle.

REFERENCE

- [1] Irina Rish 2001. An Empirical study on Naïve Bayes Classifier, Research Gate Publication -228845263
- [2] Irene Irawaty, Rachmadita Andreswari, Dita Pramesti 2019. Vectorizer Comparison for Sentiment Analysis on Social Media YouTube: A Case Study, IC2IE50715.2020.9274650