**Index**

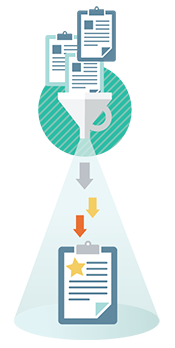
1. Introduction
2. Proposed work
3. Objectives
4. Methodology
5. Technology and software requirements

**Introduction**

The projects aims at developing a machine learning model that can evaluate resumes and by the means of resume it is implied the text in resume to categorise and filter the contents of that resume to in such a way that thousands of resumes can be analysed and filtered to specific categories and classifications which is aimed at helping faster processing of resumes according to the needs to an employer. The project mainly uses a large number of data set of resumes for its training purposes and thus after some amount of training data it can very well predict desired outcome. For UI purposes the project will be an web application which has different options to choose from for the analytics.

This project will have it front-end developed in HTML5, CSS3, JavaScript, Bootstrap and JQuery, and for the back-end Python and its framework Django will be used.

**Proposed Work**

The development of project is divided into few steps to ensure that data is formally collected and displayed in the desired format and is predicted to the best accurate levels.

The stages of project will be as follows:

* Development of UI for the web application
* Development of server side databases and data warehousing algorithms
* Implementation of data analysis factors and algorithms which will start to accept data from UI, which will then be analysed
* Analysis of gathered data by cron jobs at day to day job for a better data prediction models
* After the training data set has been feed to the system the system will be ready to accept test data and predict the desired output

**Objectives**

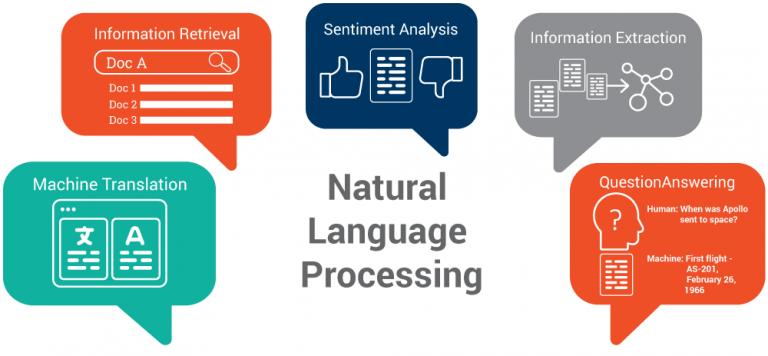
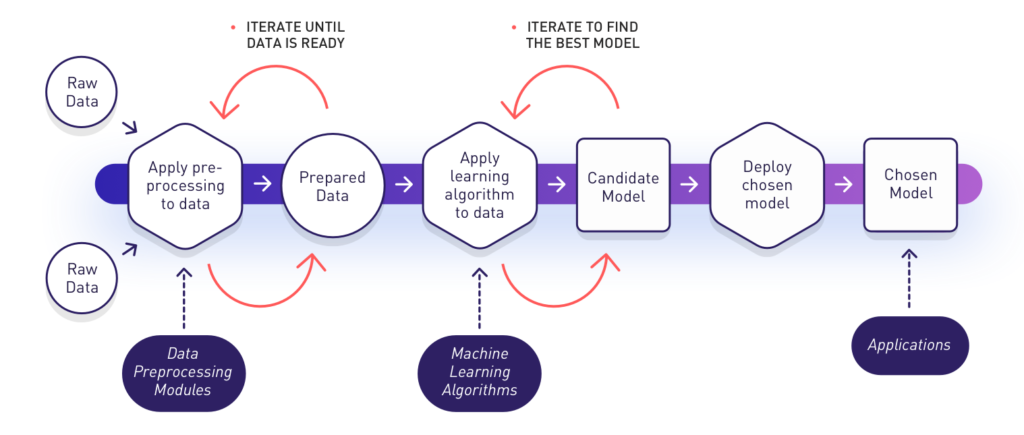
Main objectives of the project are:

* Facilitating employers to find the best suited resume for the respective jobs.
* Empowering students and grads to find what best they can perform in their respective field of studies and experience.
* Allowing employer to filter resumes based on the filters he/she wants to apply to the candidates to be chosen
* Powerful tool for big organisations where a lot of candidates apply and thus saving human efforts

**Methedology**

**Machine Learning**

Machine learning is a field of [computer science](https://en.wikipedia.org/wiki/Computer_science) that uses statistical techniques to give [computer systems](https://en.wikipedia.org/wiki/Computer_systems) the ability to "learn" (e.g., progressively improve performance on a specific task) with [data](https://en.wikipedia.org/wiki/Data), without being explicitly programmed.

The name machine learning was coined in 1959 by [Arthur Samuel](https://en.wikipedia.org/wiki/Arthur_Samuel). Machine learning explores the study and construction of [algorithms](https://en.wikipedia.org/wiki/Algorithm) that can learn from and make predictions on [data](https://en.wikipedia.org/wiki/Data) – such algorithms overcome following strictly static [program instructions](https://en.wikipedia.org/wiki/Computer_program) by making data-driven predictions or decisions, through building a [model](https://en.wikipedia.org/wiki/Mathematical_model) from sample inputs. Machine learning is employed in a range of computing tasks where designing and programming explicit algorithms with good performance is difficult or infeasible; example applications include [email filtering](https://en.wikipedia.org/wiki/Email_filtering), detection of network intruders, and [computer vision](https://en.wikipedia.org/wiki/Computer_vision).

**NLP**

Natural language processing is an area of [computer science](https://en.wikipedia.org/wiki/Computer_science) and [artificial intelligence](https://en.wikipedia.org/wiki/Artificial_intelligence) concerned with the interactions between computers and human languages, in particular how to program computers to process and analyze large amounts of [natural language](https://en.wikipedia.org/wiki/Natural_language) data.

**Technology and Software Requirements**

**HTML5**

HTML 5 is a [markup language](https://en.wikipedia.org/wiki/Markup_language) used for structuring and presenting content on the [World Wide Web](https://en.wikipedia.org/wiki/World_Wide_Web). It is the fifth and current major version of the [HTML](https://en.wikipedia.org/wiki/HTML) standard, and subsumes [XHTML](https://en.wikipedia.org/wiki/XHTML). It currently exists in two standardized forms, intended primarily for Web content developers and browser developers.

**CSS3**

Cascading Style Sheets (CSS) is a [style sheet language](https://en.wikipedia.org/wiki/Style_sheet_language) used for describing the [presentation](https://en.wikipedia.org/wiki/Presentation_semantics) of a document written in a [markup language](https://en.wikipedia.org/wiki/Markup_language) like [HTML](https://en.wikipedia.org/wiki/HTML). CSS is a cornerstone technology of the [World Wide Web](https://en.wikipedia.org/wiki/World_Wide_Web), alongside HTML and [JavaScript](https://en.wikipedia.org/wiki/JavaScript). CSS is designed to enable the separation of presentation and content, including [layout](https://en.wikipedia.org/wiki/Page_layout), [colors](https://en.wikipedia.org/wiki/Color), and [fonts](https://en.wikipedia.org/wiki/Typeface). This separation can improve content [accessibility](https://en.wikipedia.org/wiki/Accessibility), provide more flexibility and control in the specification of presentation characteristics, enable multiple [web pages](https://en.wikipedia.org/wiki/Web_page) to share formatting by specifying the relevant CSS in a separate .css file, and reduce complexity and repetition in the structural content.

**JavaScript**

JavaScript often abbreviated as JS, is a [high-level](https://en.wikipedia.org/wiki/High-level_programming_language), [interpreted](https://en.wikipedia.org/wiki/Interpreted_language) [programming language](https://en.wikipedia.org/wiki/Programming_language). It is a language which is also characterized as [dynamic](https://en.wikipedia.org/wiki/Dynamic_programming_language), [weakly typed](https://en.wikipedia.org/wiki/Weak_typing), [prototype-based](https://en.wikipedia.org/wiki/Prototype-based_programming) and [multi-paradigm](https://en.wikipedia.org/wiki/Multi-paradigm_programming_language).

Alongside [HTML](https://en.wikipedia.org/wiki/HTML) and [CSS](https://en.wikipedia.org/wiki/CSS), JavaScript is one of the three core technologies of the [World Wide Web](https://en.wikipedia.org/wiki/World_Wide_Web). JavaScript enables interactive [web pages](https://en.wikipedia.org/wiki/Web_page) and thus is an essential part of [web applications](https://en.wikipedia.org/wiki/Web_application). The vast majority of [websites](https://en.wikipedia.org/wiki/Website) use it, and all major [web browsers](https://en.wikipedia.org/wiki/Web_browser) have a dedicated [JavaScript engine](https://en.wikipedia.org/wiki/JavaScript_engine) to execute it.

**jQuery**

jQuery is a [cross-platform](https://en.wikipedia.org/wiki/Cross-platform) [JavaScript library](https://en.wikipedia.org/wiki/JavaScript_library) designed to simplify the [client-side scripting](https://en.wikipedia.org/wiki/Client-side_scripting) of [HTML](https://en.wikipedia.org/wiki/HTML). It is [free, open-source software](https://en.wikipedia.org/wiki/Free_and_open_source_software) using the permissive [MIT License](https://en.wikipedia.org/wiki/MIT_License). [Web](https://en.wikipedia.org/wiki/World_Wide_Web) analysis indicates that it is the most widely deployed JavaScript library by a large margin.

**Bootstrap**

Bootstrap is a [free and open-source](https://en.wikipedia.org/wiki/Free_and_open-source_software) front-end [framework](https://en.wikipedia.org/wiki/Application_framework) for designing [websites](https://en.wikipedia.org/wiki/Website) and [web applications](https://en.wikipedia.org/wiki/Web_application). It contains [HTML](https://en.wikipedia.org/wiki/HTML)- and [CSS](https://en.wikipedia.org/wiki/CSS)-based design templates for [typography](https://en.wikipedia.org/wiki/Typography), forms, buttons, navigation and other interface components, as well as optional [JavaScript](https://en.wikipedia.org/wiki/JavaScript) extensions. Unlike many earlier web frameworks, it concerns itself with [front-end development](https://en.wikipedia.org/wiki/Front-end_web_development) only.

**Python**

Python is an [interpreted](https://en.wikipedia.org/wiki/Interpreted_language) [high-level programming language](https://en.wikipedia.org/wiki/High-level_programming_language) for [general-purpose programming](https://en.wikipedia.org/wiki/General-purpose_programming_language). Created by [Guido van Rossum](https://en.wikipedia.org/wiki/Guido_van_Rossum) and first released in 1991, Python has a design philosophy that emphasizes [code readability](https://en.wikipedia.org/wiki/Code_readability), notably using [significant whitespace](https://en.wikipedia.org/wiki/Significant_whitespace). It provides constructs that enable clear programming on both small and large scales. In July 2018, Van Rossum stepped down as the leader in the language community after 30 years.

Python features a [dynamic type](https://en.wikipedia.org/wiki/Dynamic_type) system and automatic [memory management](https://en.wikipedia.org/wiki/Memory_management). It supports multiple [programming paradigms](https://en.wikipedia.org/wiki/Programming_paradigm), including [object-oriented](https://en.wikipedia.org/wiki/Object-oriented_programming), [imperative](https://en.wikipedia.org/wiki/Imperative_programming), [functional](https://en.wikipedia.org/wiki/Functional_programming) and [procedural](https://en.wikipedia.org/wiki/Procedural_programming), and has a large and comprehensive [standard library](https://en.wikipedia.org/wiki/Standard_library)

**Django**

Django is a [free and open-source](https://en.wikipedia.org/wiki/Free_and_open-source_software) [web framework](https://en.wikipedia.org/wiki/Web_framework), written in [Python](https://en.wikipedia.org/wiki/Python_(programming_language)), which follows the model-view-template (MVT) [architectural pattern](https://en.wikipedia.org/wiki/Architectural_pattern_(computer_science)).

Django's primary goal is to ease the creation of complex, database-driven websites. Django emphasizes [reusability](https://en.wikipedia.org/wiki/Reusability) and "pluggability" of components, less code, low coupling, rapid development, and the principle of [don't repeat yourself](https://en.wikipedia.org/wiki/Don%2527t_repeat_yourself). Python is used throughout, even for settings files and data models. Django also provides an optional administrative [create, read, update and delete](https://en.wikipedia.org/wiki/Create,_read,_update_and_delete) interface that is generated dynamically through [introspection](https://en.wikipedia.org/wiki/Type_introspection) and configured via admin models.

**Scikit-learn and NumPy**

Scikit-learn is a [free software](https://en.wikipedia.org/wiki/Free_software) [machine learning](https://en.wikipedia.org/wiki/Machine_learning) [library](https://en.wikipedia.org/wiki/Library_(computing)) for the [Python](https://en.wikipedia.org/wiki/Python_(programming_language)) programming language.[[3]](https://en.wikipedia.org/wiki/Scikit-learn%23cite_note-jmlr-3) It features various [classification](https://en.wikipedia.org/wiki/Statistical_classification), [regression](https://en.wikipedia.org/wiki/Regression_analysis) and [clustering](https://en.wikipedia.org/wiki/Cluster_analysis) algorithms including [support vector machines](https://en.wikipedia.org/wiki/Support_vector_machine), [random forests](https://en.wikipedia.org/wiki/Random_forests), [gradient boosting](https://en.wikipedia.org/wiki/Gradient_boosting), [k-means](https://en.wikipedia.org/wiki/K-means_clustering) and [DBSCAN](https://en.wikipedia.org/wiki/DBSCAN), and is designed to interoperate with the Python numerical and scientific libraries [NumPy](https://en.wikipedia.org/wiki/NumPy) and [SciPy](https://en.wikipedia.org/wiki/SciPy).