Development of

Supply Chain Management System

A Practicum Report Submitted by

**Md Shahriair Alam**

**ID: 20103204**

In the Partial Fulfillment of the Requirements

for the Award of Bachelor of Computer Science and Engineering (BCSE)



Department of Computer Science and Engineering

College of Engineering and Technology

IUBAT – International University of Business Agriculture and Technology

Fall 2023

Development of Supply Chain Management System

Md Shahriair Alam (20103204)

A Practicum Report Submitted in the Partial Fulfillment of the Requirements for the Award of Bachelor of Computer Science and Engineering (BCSE)

The thesis has been examined and approved,

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Prof Dr. Utpal Kanti Das**Chairman & Professor   
Dept. of Computer Science and Engineering  
IUBAT – International University of Business Agriculture and Technology

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Rashedul Islam**Co-supervisor, Coordinator and Assistant Professor  
Dept. of Computer Science and Engineering  
IUBAT – International University of Business Agriculture and Technology

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Tasnia Anjum Rifah**Supervisor and Lecturer  
Dept. of Computer Science and Engineering  
IUBAT – International University of Business Agriculture and Technology

Fall - 2023

**Letter of Transmittal**

19 October 2023

The Chair

Practicum and Placemen Board

Department of Computer Science and Engineering

IUBAT–International University of Business Agriculture and Technology

4 Embankment Drive Road, Sector 10, Uttara Model Town

Dhaka 1230, Bangladesh

**Subject:** Letter of Transmittal.

Dear Sir,

With due respect, I would like to inform you that it is a great pleasure and a great pleasure for me to submit this report entitled “**Development of Supply Chain Management System**” to complete my Practicum course.

It was a great opportunity for me to work on this project to make my theoretical knowledge more realistic and I gained a lot of exposure to the business culture of a famous company. I now look forward to your kind commentary on this performance report.

I will always be very grateful to you if you kindly go through this report and check my performance.

Sincerely Yours,

\_\_\_\_\_\_\_\_\_\_\_\_\_

Md Shahriair Alam

20103204

**Student’s Declaration**

I am Md Shahriair Alam, is student of the BCSE-Bachelor of Computer Science and Engineering program, under the College of Engineering and Technology (CEAT) of the International University of Business Agriculture and Technology (IUBAT) announcing this report entitled **“Development of Supply Chain Management System”** has been prepared for the completion of the P\practicum course, which is part of the Bachelor of Computer Science and engineering degree.

The report of **“Development of Supply Chain Management System”** was edited by me. All modules and procedures for this research are done after proper testing and online information.

It is not designed for other purposes, awards or presentations.

\_\_\_\_\_\_\_\_\_\_\_\_\_

Md Shahriair Alam

20103204

**Supervisor’s Certification**

This is to ensure that the practicum report on the **“Development of Supply Chain Management System”** is compiled by Md Shahriair Alam, with ID #20103204, of IUBAT– International University of Business Agriculture and Technology, as part of the fulfillment of the required part of an effective praacticum course. The report has been prepared under my supervision and is a record of the work accomplished, successfully completed. To the best of my knowledge and as per there declaration, no portions of this report have been posted anywhere by any degree, diploma or certificate.

You are now allowed to submit a report. I wish your every success in the future endeavors.

Practicum Supervisor

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tasnia Anjum Rifah

Lecturer

Department of Computer Science and Engineering

IUBAT–International University of Business Agriculture and Technology

**Department’s Certification**

On behalf of the Department of Computer Science and Engineering, IUBAT-International University of Business Agriculture and Technology, I undersigned, confirm the performance report on ‘**Development of Supply Chain Management System**’ for Bachelor of Computer Science and Engineering (BCSE) degrees was duly presented by Md Shahriair Alam (ID No: 20103204) and approved by the department.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tasnia Anjum Rifah

Lecturer

Department of Computer Science and Engineering

IUBAT- International University of Business Agriculture and Technology

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Rashedul Islam

Coordinator and Assistant Professor

Department of Computer Science and Engineering

IUBAT- International University of Business Agriculture and Technology

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Prof. Dr. Utpal Kanti Das

Chairman & Professor

Department of Computer Science and Engineering

IUBAT- International University of Business Agriculture and Technology

**Abstract**

**Acknowledgments**

Firstly, I want to give Almighty Allah to our sincere gratitude for His wonderful grace, which enabled us to successfully finish the thesis.

I would like appreciate and acknowledge all the people who has encouraged and supported me through the entire journey and has played a great role the completion of my Practicum and the report on “**Development of Supply Chain Management System**”. First of all I would like to thank Almighty for uncountable and beyond to words reasons and for giving me the power to stay motivated all the time for completing any task. I want to thank my parents, who have endowed their immeasurable support and encouragement to attain this exquisite event of my life.

I would like to convey sincere thanks to our Vice Chancellor **Dr. Abdur Rab** for giving me an opportunity to submit this report. My outmost and sincere gratitude goes to **Prof. Dr. Utpal Kanti Das**, Chairman of Department of Computer Science and Engineering and **Rashedul Islam**, Coordinator of Department of Computer Science and Engineering for allowing me to work on the project.

I would like to thank our academic supervisor, **Tasnia Anjum Rifah**, Lecturer of the Department of Computer Science and Engineering of IUBAT-International University of Business Agriculture and Technology, who has encouraged me throughout the entire time and has provided me with all the necessary guidelines and support that I needed.

At last, I would like to thank Tohidul Islam (CTO, Kodeeo Limited) for giving me the opportunity to complete my internship and project at Kodeeo Limited.

**INTERNSHIP CERTIFICATE**

**Chapter I. About Organization**

1.1 Organization Overview

Kodeeo Limited is a Software development company at Bangladesh that provides IT consulting and services in the fields of software development, website and mobile application design and development and the implementation of e-commerce. Kodeeo Limited has developed significant competence over the past four years and has successfully completed projects of all sizes while providing solutions to end users, businesses, organizations, associations, and governmental entities.

Kodeeo Limited, based on client‘s requirement, puts its professional team and process planning to work and is able to attain goals and objectives for the benefit of the client. These are based on a win-win mindset, commitment and immense planning to benefit customers directly. Our professionals have several years of experience, providing focused and targeted client- oriented solutions to our customers.

1.2 Organizational Services

Kodeeo Limited provides various types services to its clients. The main services provided by Kodeeo Limited are,

* Software Development
* Mobile App Development
* Website Development
* E-commerce Website
* ERP
* Domain and Hosting

1.2.1 Software Development

A functional software or web application benefits businesses of all sizes to run better through digital transformation and modern capabilities. Kodeeo Limited plans, designs, develops, tests, deploys and maintains software and web applications to drive efficiency and effectiveness.

For building efficient and client satisfactory softwares Kodeeo Limited has the followings,

* Cloud-based and On-site custom software application development, integration, migration, and maintenance Development of connected mobile applications on iOS, and Android platforms using native and cross-platform technologies.
* Certified and trained engineers on well-known clouds such as AWS
* DevOps workflow automation, continuous integration and deployment pipeline, microservices architecture and dockerization
* Automated and manual software testing services for software and products

Kodeeo Limited has been designing, architecting & delivering technology and business solutions to its clients.

1.2.2 Website Development

A website is a company’s Online Identity. Websites can boost awareness of a brand or product, leading to more sales or brand awareness depending on the company‘s target. A website presents a company to the world, builds relationships with customers and target clientele.

Kodeeo Limited can build for an effective website that will get more traffic from the search engines, fast-loading web pages with excellent navigation and designing, to communicate more effectively across the Globe. Whether companies are looking for a Dynamic site or static site is a solution Kodeeo Limited can help acquire either.

Depending on the client‘s requirement, Kodeeo Limited can offer the most effective solution through websites made to target potential customers and make company‘s presence live on web.

1.2.3 Mobile App Developement

In order to make a Mobile App to be successful, the objectives need to be assessed and it is needed to see the big picture. Kodeeo Limited connects its customers to the world by designing and developing conversational Native Mobile Applications that function perfectly on mobile devices run on operating systems like Android, iOS, and Windows.

Kodeeo Limited also has exceptional expertise in scaling an application‘s user bases and developing growth strategies through advance machine learning processes that focus on mostly user behaviour in essence of providing valuable insights for optimization and improvement.

Kodeeo Limited has worked with leading B2B Networking Platforms, Live Tracking Applications, Delivery Systems, Event Registration Portals, and so much more. Kodeeo Limited has created high-end mobile apps with the help of a team consisting of individuals with more than a decade of experience.

1.2.4 Domain and Hosting

A domain name is your name on the internet. Imagine you have a phone agenda, where you mark each person’s name and then its phone number. You won’t probably remember all the phone numbers of everyone, right?

So a domain name is like a world wide web agenda. Whenever you type the URL or domain name of a company, your browser will query the DNS Server of the hosting provider, and then it will obtain the website’s IP address. With the IP address in hands, your browser will reach the desired website.

In other words, the IP address is like the phone number of your website. But, it’s identified by the domain name, instead of a personal name in your phone’s agenda.

Kodeeo limited provides best cheap rate domain and hosing facility for the clients.

1.3 Clients of the Organization

1.4 Organizations Expertise and Goals

Kodeeo Limited authenticates their goals clearly on each of these crucial parameters and follow processes and standards to measure achievement continually and stay on target. There is no other way to go about it, is what each individual believes. Kodeeo Limited‘s high score on the customer appreciation indicator is a measure of how consistently they are able to deliver on these crucial parameters.

Having operations in United States and South Asia, a strong presence and full understanding of today‘s business dynamics gives Kodeeo Limited the ability to strengthen existing relationships, build new ones & enter new markets. On the other hand, Kodeeo Limited‘s alliance partners give them an extended reach into markets where they do not have a direct presence. Kodeeo Limited can assure their clients to be the One-Stop partner in their line of expertise. Kodeeo Limited‘s Vision has been and always will be to be a part of Client‘s success.

1.5 Organizational Structure

**Chapter 2. Project Introduction**

**2.1 Introduction**

In the ever-evolving landscape of the bakery industry, where precision and efficiency are paramount, the adoption of a well-designed Supply Chain Management (SCM) system emerges as a strategic necessity. Bakeries, facing the dual challenges of fluctuating ingredient costs and the need to stay agile in response to dynamic market trends, find themselves at a crossroads. Conventional methods, often manual or lacking in organization, can result in operational inefficiencies and increased financial burdens.

This project report is dedicated to exploring the imperative of introducing a tailored SCM system in bakery operations. The primary focus is to optimize processes seamlessly, encompassing procurement, production, and distribution, ultimately leading to improved operational efficiency and the ability to meet the heightened demands of the market

The report will thoroughly examine the project's objectives, delving into the key components of the proposed SCM system, exploring opportunities for technology integration, and projecting the anticipated benefits for bakery operations. By simplifying complex processes and harnessing the power of practical technology solutions, this project aspires to empower bakeries, enabling them to navigate challenges adeptly and meet consumer expectations with ease in the competitive bakery market. Throughout the report, emphasis will be placed on the simplicity and practicality of the proposed SCM system to make it accessible and beneficial for bakery stakeholders at all levels.

2.2 Project Overview

This project aims to introduce a tailored Supply Chain Management (SCM) system to optimize bakery operations. Faced with challenges such as fluctuating ingredient costs and the need for quick adaptation to market trends, the bakery industry requires a streamlined approach. The objectives include improving procurement, production planning, inventory management, supplier relationships, and distribution. Key components of the SCM system encompass inventory management, supplier relationship management (SRM), production planning and scheduling, and distribution and logistics. Leveraging modern technologies like IoT, RFID, and AI, the project seeks to enhance real-time tracking and decision-making. Anticipated benefits include enhanced visibility, cost optimization, and improved customer satisfaction. The project's success will be measured through key milestones outlined in a comprehensive timeline, with the ultimate goal of contributing to the sustained success of bakery enterprises in a competitive market.

2.3 Objects

The main objective of the Project on Supply Chain Management System is to manage the details of Customer, Product, Product Company, Order, Shipment. It manages all the information about Customer, Delivery, Shipment, Customer. The project is totally built at administrative end and thus only the administrator is guaranteed the access. The purpose of the project is to build an application program to reduce the manual work for managing the Customer, Product, Delivery, Product Company. It tracks all the details about the Product Company, Order, Shipment.

The primary objectives of this project include:

* Streamlining procurement processes to ensure timely and cost-effective sourcing of raw materials.
* Optimizing production planning and scheduling for improved resource utilization and reduced lead times.
* Implementing an efficient inventory management system to maintain optimal stock levels and minimize waste.
* Strengthening supplier relationships through a collaborative Supplier Relationship Management (SRM) platform.
* Enhancing distribution and logistics to ensure timely and cost-efficient product delivery.

2.4 Scope of the project

It may help collecting perfect management in details. In a very short time, the collection will be obvious, simple and sensible. It will help a person to know the management of passed year perfectly and vividly. It also helps in current all works relative to Supply Chain Management System. It will be also reduced the cost of collecting the management & collection procedure will go on smoothly.

Our project aims at Business process automation, i.e. we have tried to computerize various processes of Supply Chain Management System.

* In computer system the person has to fill the various forms & number of copies of the forms can be easily generated at a time.
* In computer system, it is not necessary to create the manifest but we can directly print it, which saves our time.
* To assist the staff in capturing the effort spent on their respective working areas.
* To utilize resources in an efficient manner by increasing their productivity through
* automation.
* The system generates types of information The system generates types of information that can be used for various
* purposes.
* It satisfies the user requirement
* Be easy to understand by the user and operator
* Be easy to operate
* Have a good user interface
* Be expandable
* Delivered on schedule within the budget. that can be used for various purposes.

2.5 Methodology

The creation of the "Supply Chain Management System" has followed an incremental model, aligning with the structural framework outlined in Software Analysis and Design. This preliminary study on the development of the Supply Chain Management System emphasizes its early stages, with a focus on establishing a foundational understanding of the system's structure and functionality. The incremental approach adopted in the development process ensures a systematic and iterative enhancement of the system, aligning with best practices in software engineering for a comprehensive and effective Supply Chain Management System.

2.6 Limitation of the project

As mentioned earlier in this report, an internship serves as a bridge connecting theoretical knowledge to practical application. The internship program at IUBAT has provided me with a valuable opportunity to witness the translation of theory into real-world practice. Initially, upon joining the organization, I encountered numerous conditions and systems that were unfamiliar to me. Over time, however, I have become accustomed to these terms and have learned from the experiences of older classmates and friends who have undergone similar adjustments. It's important to note that a limitation of this internship project is the constraint of time. Developing comprehensive software within a short timeframe is a challenging task, even with the application of software engineering techniques. Consequently, the scope of the internship project had to be narrowed due to these time constraints.

2.7 Process model

I chose incremental process model for the development of the system. The incremental model is a developed version of the waterfall model. This product is designed, implemented, integrated and tested as a series of incremental builds.The reasons for which I have selected the Incremental Process model are as follows,

The incremental model prioritizes requirements of the system and then implements

them in groups.

* Develop high-risk or major functions first
* Each release delivers an operational product
* Customer can respond to each build
* Lowers initial delivery cost
* Initial product delivery is faster
* Customers get important functionality early

**Chapter 3. Requirement Engineering**

3.1 Requirement Analysis

The requirement analysis serves as a tool for the software designer to convert data, architectural, interface, and component level designs into information, function, and behavior. The completion of requirement analysis occurred during the task phases outlined in this chapter.

3.2 Requirement Engineering

Requirements engineering, as implied by its name, is a branch within engineering dedicated to the development of software systems and the identification of user needs. While various definitions exist for requirements engineering, they all converge on the idea that it revolves around comprehending user expectations for a computer system and deciphering the implications of these requirements on the design. This field shares close ties with software engineering, which principally concentrates on the process of constructing systems aligned with customer preferences.

The following are the results of the requirement engineering for this project:

* User requirements
* System requirements
* Functional requirements
* Non-Functional requirements

3.2.1 User Requirements

The Supply Chain Management System has 6 types of users,

1. Admin
2. Supplier
3. Manufacturer
4. Retailer
5. Distributor
6. Customer

The user requirements of the project Supply Chain Management System are as follows:

**Chapter III. Research Methodology**

3.1 Recap of the research questions

The research question focuses on the development of an intelligent career guidance system that utilizes machine learning algorithms to provide accurate and personalized career advice to users. This research question aims to investigate the various machine learning algorithms that can be used to analyze the user's personal and professional characteristics, identify their skills, strengths, weaknesses, and interests, and suggest suitable career paths based on their profile. This question requires exploring the existing literature on the various machine-learning techniques and determining which algorithms are most effective in predicting and recommending career options. Then another research question focuses on the design and development of a user-friendly interface for the intelligent career guidance system. This question requires investigating the various design principles and guidelines for developing a user interface that is intuitive, easy to use, and aesthetically pleasing. The research will also investigate the usability of the system and gather feedback from users to improve the interface and overall user experience.

Another research question aims to evaluate the effectiveness of the intelligent career guidance system in terms of its accuracy, reliability, and usefulness. This research question requires conducting a series of tests to measure the system's ability to accurately predict and recommend suitable career paths for users. The research will also investigate the reliability of the system by comparing its recommendations to those of human career advisors. The usefulness of the system will be evaluated by gathering feedback from users on the quality and relevance of the career advice provided by the system.The next research question focuses on the ethical implications of using machine learning algorithms in career guidance. This research question requires investigating the potential ethical concerns that may arise from using machine learning algorithms to predict and recommend career paths. The research will investigate issues such as bias, privacy, and transparency and develop ethical guidelines for the use of machine learning algorithms in career guidance.

Overall, the research questions in this thesis topic aim to investigate the development of an intelligent career guidance system using machine learning algorithms, focusing on the design and development of a user-friendly interface, evaluating the effectiveness of the system, and investigating the ethical implications of using machine learning algorithms in career guidance. By addressing these research questions, the thesis will contribute to the development of a more accurate, personalized, and ethical approach to career guidance, which can help individuals make informed decisions about their careers and achieve their professional goals.

3.2 Description of the method

Selecting the right algorithm for implementing machine learning can be a daunting task. It's essential to preprocess data before using algorithms to avoid bad results. In a study, two algorithms were used to recommend careers for students, and the k-nearest neighbor was found to be the most accurate. Data preparation and preprocessing are crucial steps in the proposed methodology. After preprocessing, algorithms were applied, and the results were compared to determine the best algorithm. Therefore, selecting the right algorithm and preprocessing data can significantly improve the accuracy of machine learning outcomes.

In figure 3.2.1 the proper steps of the proposed methodology are encapsulated below:

Dataset

Label Encoding

Data Splitting

Data Preprocessing

Trained model

Training Data

Testing Data

Training the model

Performance evaluation

* Filling Missing Data
* Correct Spelling Mistake
* Remove Stop Word

Figure: 3.2.1 Proposed methodology of work

FIGURE: 3.2.1 PROPOSED METHODOLOGY OF WORK

FIGURE: 3.2.1 PROPOSED METHODOLOGY OF WO RK

Figure 13.2.1 Proposed methodology of work

3.3 Dataset Description

Due to difficulties with collecting datasets, we resorted to using dummy datasets that were specifically created from a Bangladeshi career perspective. Although dummy data does not contain any valuable information, it can be used to reserve space for real data and for both operational and testing purposes. We used dummy datasets for testing purposes by creating questions related to students' interests, academic backgrounds, and environmental situation. By gathering the answers to these questions, we can predict the appropriate field of study for a specific student. Various fields of study were included so that students could choose their own area of study by answering questions to a machine, which could also determine their future job opportunities based on their interest’s chart of the data reveals that we included ten distinct areas of study in our dummy dataset to obtain the recommended output for students. We used ten categories to suggest the best field of study for each student.

In figure 3.3.1 the illustration of our dataset has been given below:

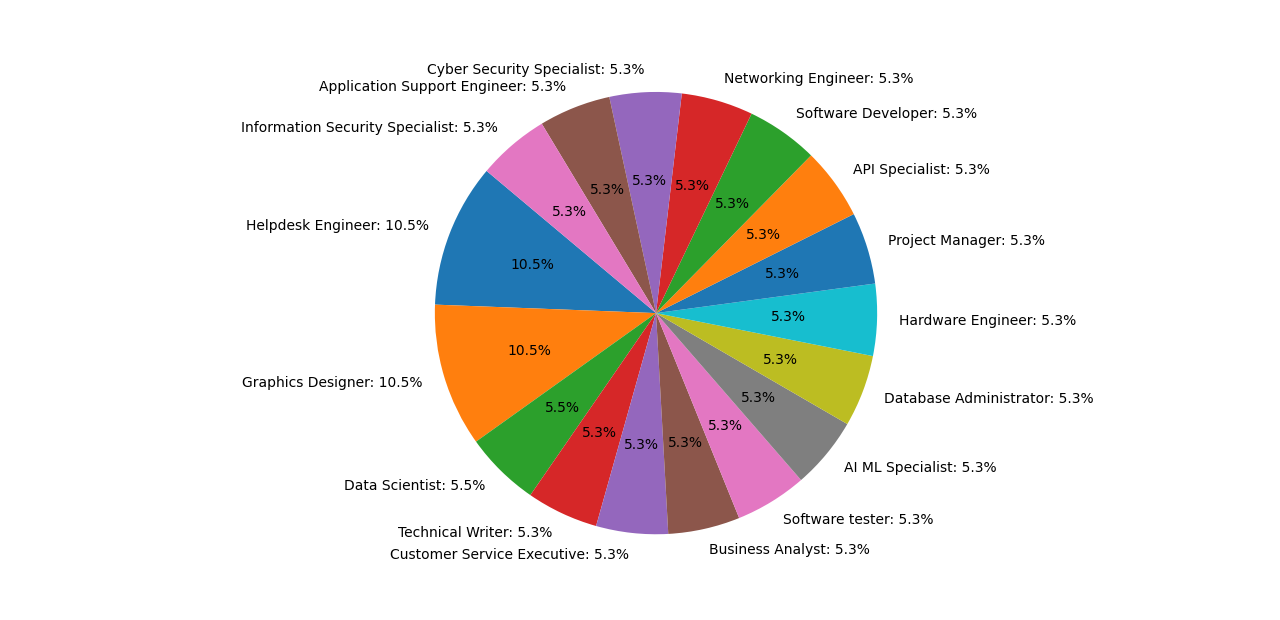


Figure 3.3.1 The illustration of our dataset

#### Data Statistics

Statistics of my Dataset –

TABLE 3.3.1: DATA STATISTICS

|  |  |
| --- | --- |
| Names | Numbers |
| Question of skillsets | 17 |
| Categories | 17 |
| Students | 100000 |

The table indicates that the questions and categories of skillsets pertain specifically to computer skills and that there are more than 1000 students involved in this scenario. This suggests that the questions and categories of skillsets are likely related to computer programming, IT, or another field of technology and that the students in question are likely studying or training in this area.

3.4 Data preprocessing

Data preprocessing in machine learning involves a series of essential steps to prepare raw data for model training and evaluation. This includes addressing missing values, outliers, and errors in the dataset, transforming categorical variables into numerical formats, scaling features for uniformity, splitting data into training, validation, and test sets, and potentially reducing dimensionality for efficiency. Proper data preprocessing ensures that the data is cleaned, organized, and structured in a way that facilitates the training of machine learning models, leading to more accurate and reliable results when making predictions or drawing insights from the data.

#### 3.4.1 Handling Missing Data

Missing data is a common issue in datasets, and it can result in biased or inaccurate results if not handled properly. This preprocessing technique involves dealing with missing values in a dataset. Several strategies for handling missing data include removing rows or columns with missing values, imputing missing values with statistical measures (such as mean, median, or mode), or using more advanced techniques like interpolation or predictive modeling to estimate missing values.

##### Mode Imputation Technique:

Mode imputation is a simple and commonly used method for handling missing data in categorical features. It helps to ensure that your dataset remains complete, which is important for training machine-learning models. However, it may not always be the best approach, as it can introduce bias if the missing data is not missing completely at random or if the mode is not a representative value for the missing cases. Other imputation techniques like mean imputation for numerical variables or more advanced methods like K-nearest neighbor’s imputation can be considered depending on the nature of your data and the problem you are trying to solve.

|  |
| --- |
| **Database Fundamentals** |
| Professional |
| Professional |
| Average |
| Poor |
|  |
| Beginner |

|  |
| --- |
| **Database Fundamentals** |
| Professional |
| Professional |
| Average |
| Poor |
| Professional |
| Beginner |

Figure 3.4.1.1 The illustration of missing data

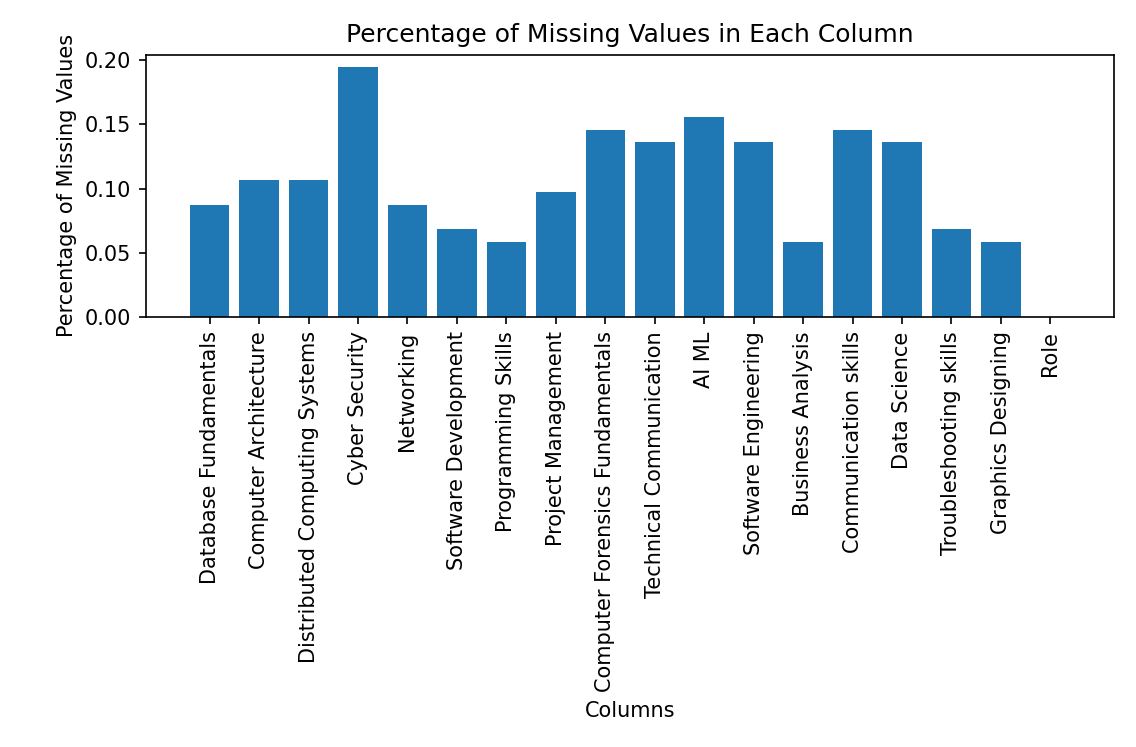


Figure 3.4.1.2 Percentage of missing value in each column

After applying the mode imputation technique, the data set column look like-

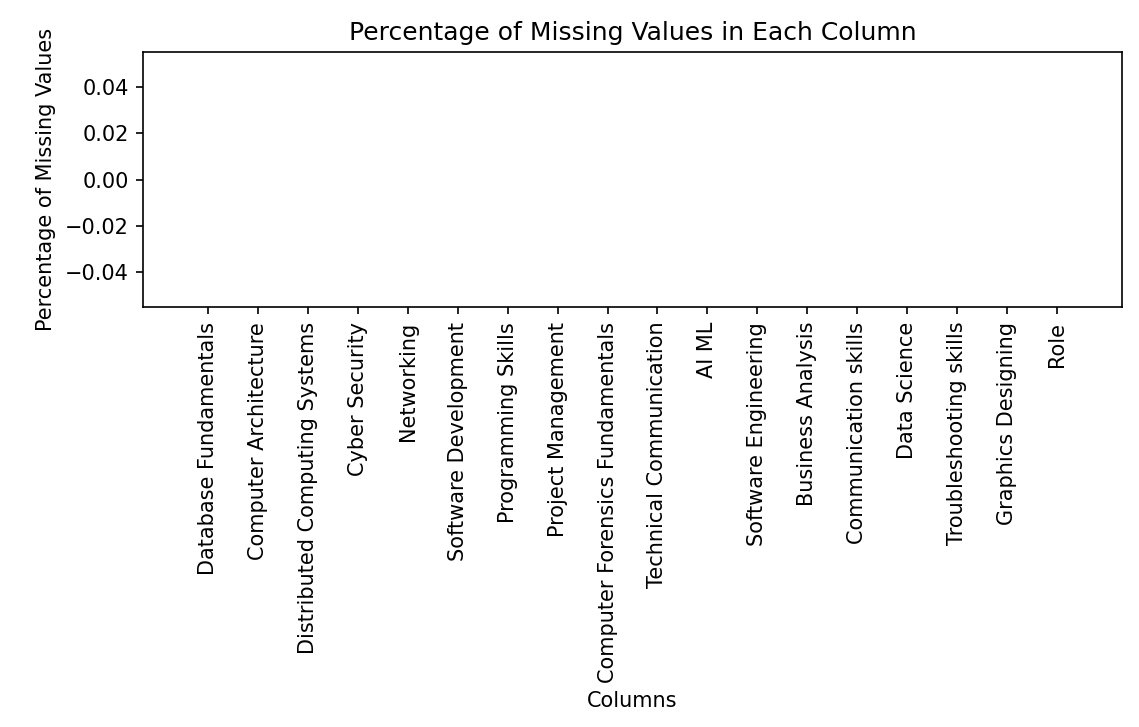


Figure 3.4.1.3 Percentage of missing value in each column

#### 3.4.2 Correct Spelling Mistake

Correcting spelling mistakes in machine learning typically involves using natural language processing (NLP) techniques. Correcting spelling mistakes in machine learning is a common NLP task and is vital for improving the user experience in various applications involving text input and processing. It's worth noting that modern spell checkers often use more advanced techniques, including deep learning models like transformers, to achieve high accuracy in correcting spelling mistakes. It employs a dictionary to define valid categories for different fields and utilizes fuzzy string matching to find and replace categories with their closest valid counterparts, ensuring data consistency. This approach is commonly used to clean and correct categorical data before further analysis or modeling.

|  |
| --- |
| **Database Fundamentals** |
| Professional |
| Professional |
| Averag |
| Poor |
| rofessional |
| Beginner |

|  |
| --- |
| **Database Fundamentals** |
| Professional |
| Professional |
| Average |
| Poor |
| Professional |
| Beginner |

Figure 3.4.2.1 The illustration of spelling mistake

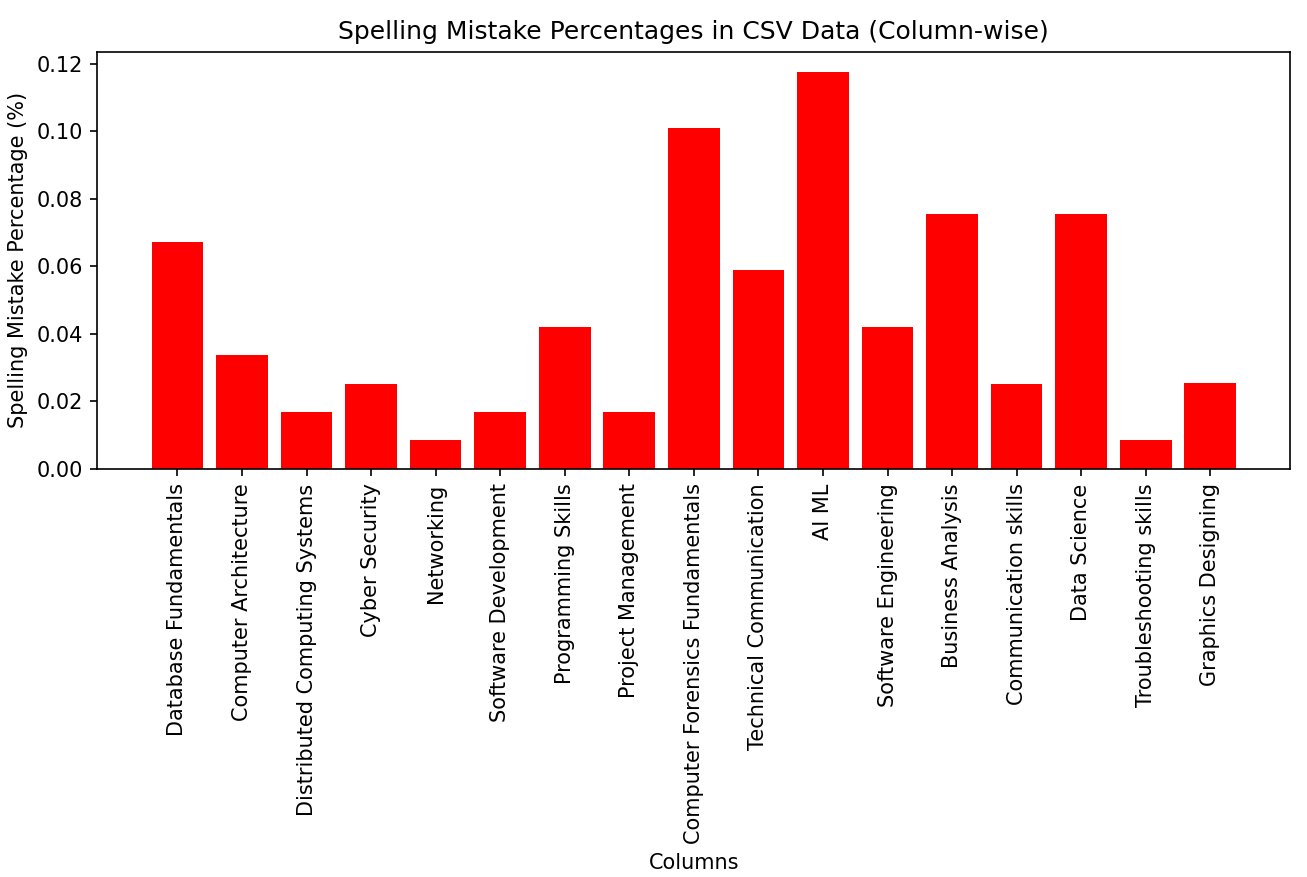


Figure 3.4.2.2 Showing percentage of mistake for each column

After applying fuzzy string matching technique for each column, we have removed the spelling mistake. The updated dataset spelling missing percentage look like-

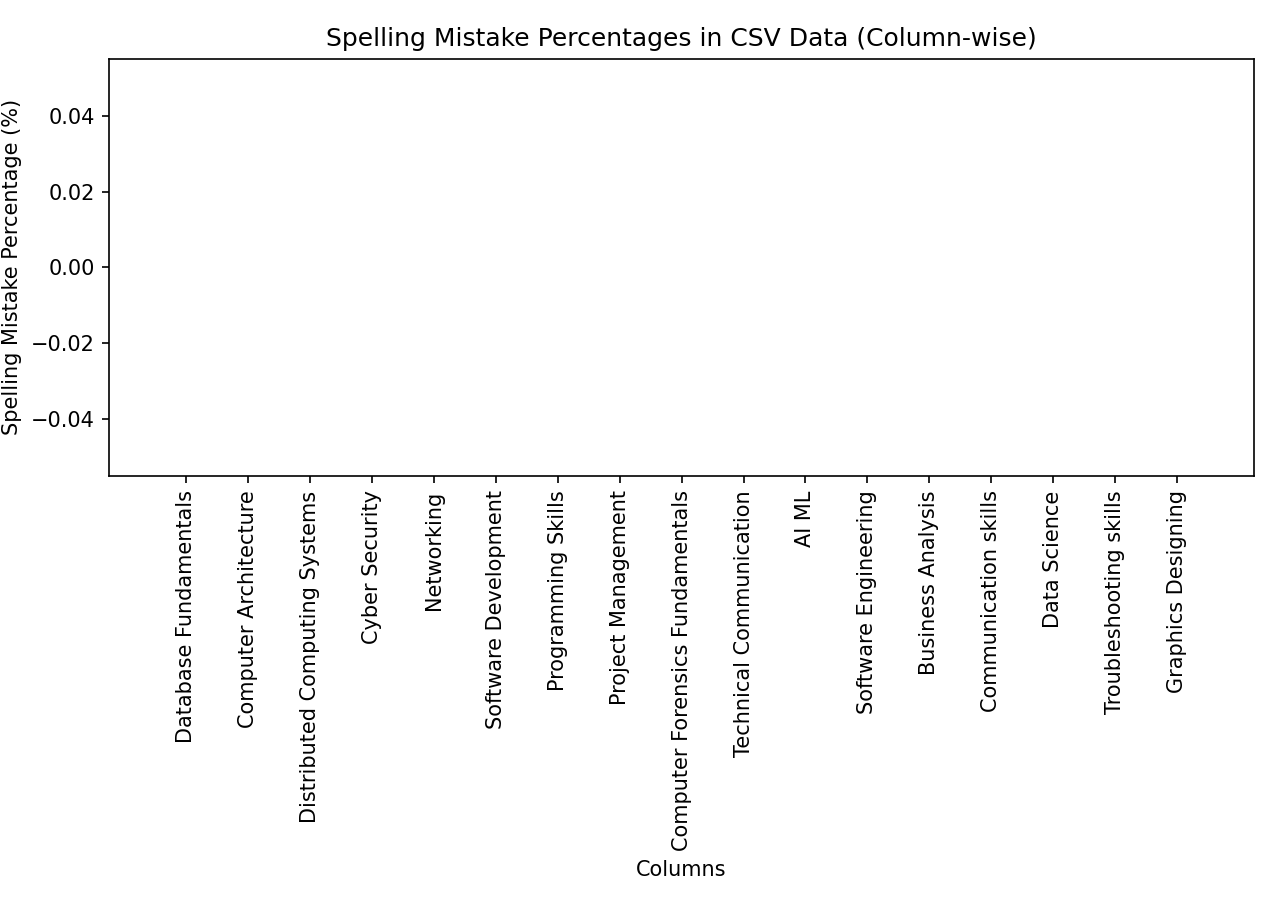


Figure 3.4.2.3 Showing percentage of mistake for each column

#### 3.4.3 Remove Stop word

Removing stop words is a standard text preprocessing method in natural language processing (NLP) that involves eliminating common words, such as "the," "is," "and," and "in," from a text. These words are considered low informational value because they appear frequently in most textbooks and don't carry specific meanings. By removing stop words, NLP models can focus on more meaningful text content, leading to better outcomes in tasks like sentiment analysis, text classification, or information retrieval. It's a fundamental step to reduce noise and improve the efficiency of text analysis.

|  |
| --- |
| **Database Fundamentals** |
| Professional |
| Professional is |
| Average |
| Poor |
| Professional i |
| Beginner on |

|  |
| --- |
| **Database Fundamentals** |
| Professional |
| Professional |
| Average |
| Poor |
| Professional |
| Beginner |

Figure 3.4.3.1 The illustration of Stop word process

Before removing the stop word from our dataset the percentage of the stop of our dataset -

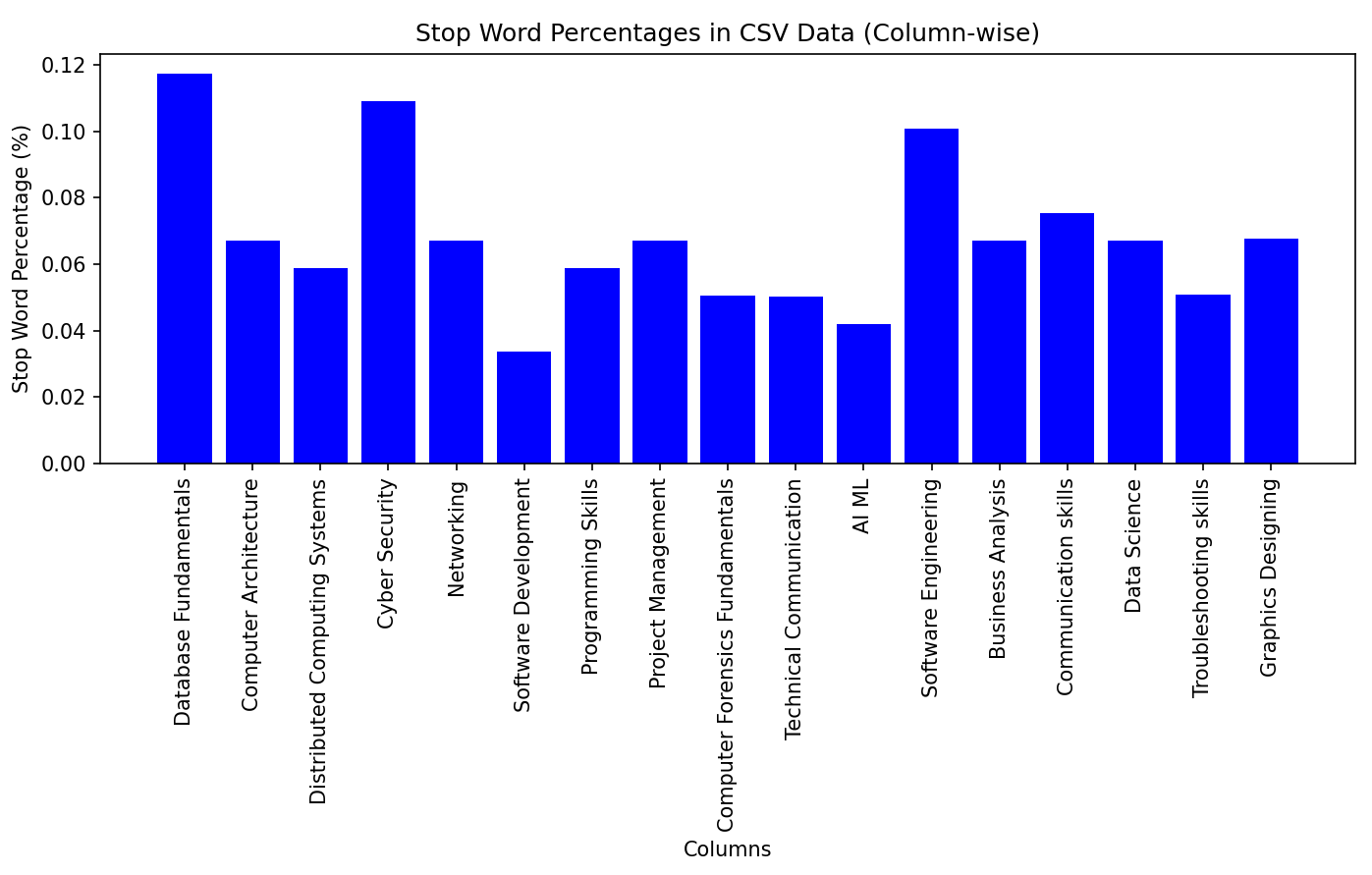


Figure 3.4.3.2 Showing percentage of stop for each column

Apply NLTK by importing it, preprocessing text, tokenizing (splitting) it, optionally removing common words (stop words), and conducting text analysis or machine learning tasks as needed, with iterative refinement and proper documentation. After applying-

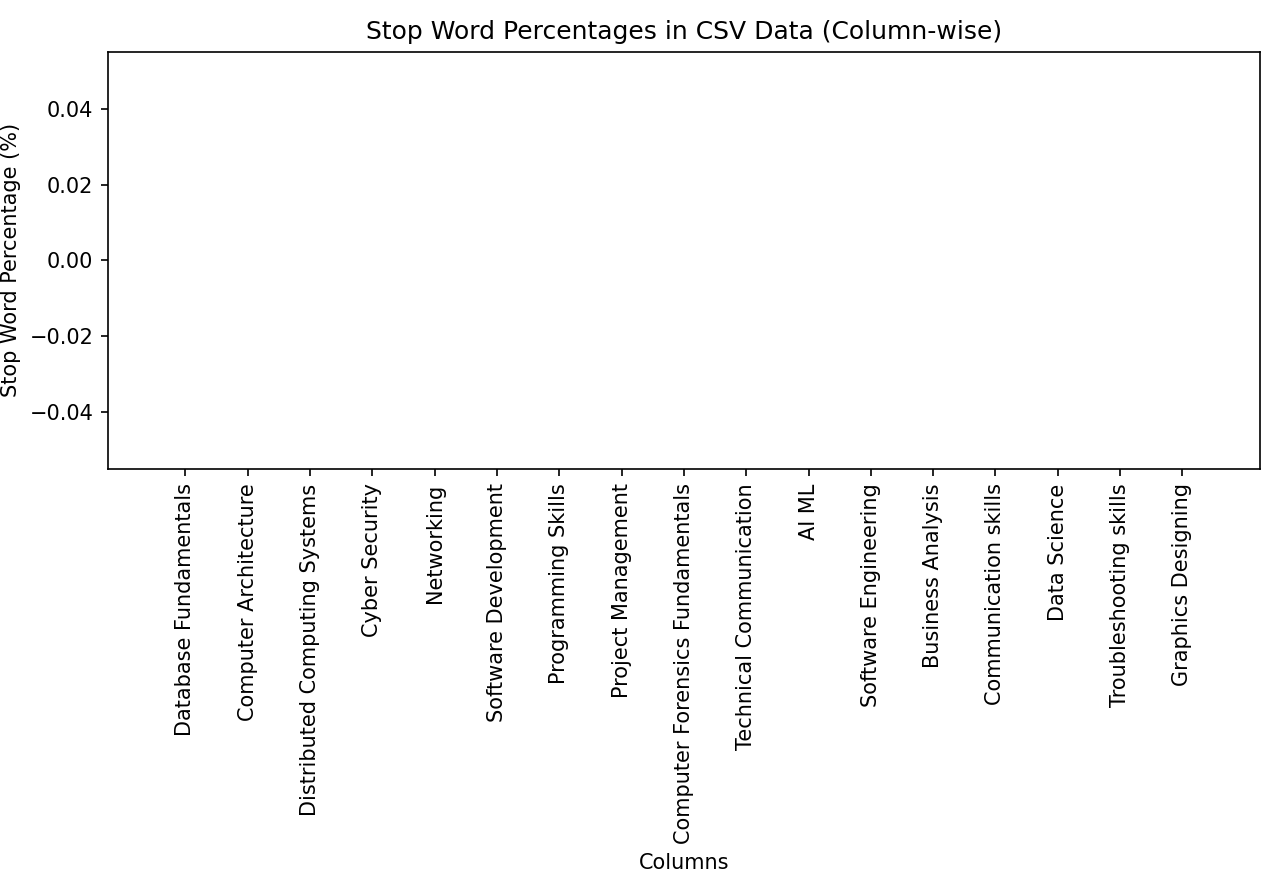


Figure 3.4.3.3 Percentage after removing stop word

3.5 Machine Learning Algorithms

The way in which an AI system performs its task, typically making predictions of output values based on input data, is referred to as a machine-learning algorithm.

#### 3.5.1 Naïve Bayes

Naïve Bayes is a classification method that relies on Bayes' Theorem, assuming that the predictors are independent. Essentially, the Naive Bayes classifier assumes that the existence of a particular feature in a category is not connected to the existence of any other feature. The Naive Bayes algorithm is adaptable for various applications, including real-time predictions, multi-class predictions, text classification, sentiment analysis, and recommendation systems.

Bayes' theorem offers a method for computing the posterior probability P(A|B) using P(A), P(B), and P(B|A).

The formula:

Where -

P(A|B) = Posterior Probability of the Hypothesis given that the Evidence is True

P(B|A) = Likelihood of the Evidence given that the Hypothesis is True

P(A)= Prior Probability of the Hypothesis

P(B)= Prior Probability that the evidence is True

3.5.2 K-NN algorithm

Given that we are developing a recommendation system, the k-NN algorithm is a suitable option for our dataset. The k-nearest neighbor (k-NN) algorithm is a supervised machine learning technique that can be used for both classification and regression tasks, although it is mainly used for classification in the industry. It works by calculating the distance between data points, which is done using the Euclidean Distance formula.

Formula:

The Euclidean Distance formula can be used to calculate the distance between two points in 2-dimensional space. This formula is commonly known as the Euclidean distance formula and is widely used for measuring distances in the plane.

If we consider two points (x1, y1) and (x2, y2), the distance between them can be calculated using the Pythagorean Theorem as follows:

(x2, y2)

d

(y2-y1)

(x1, y1)

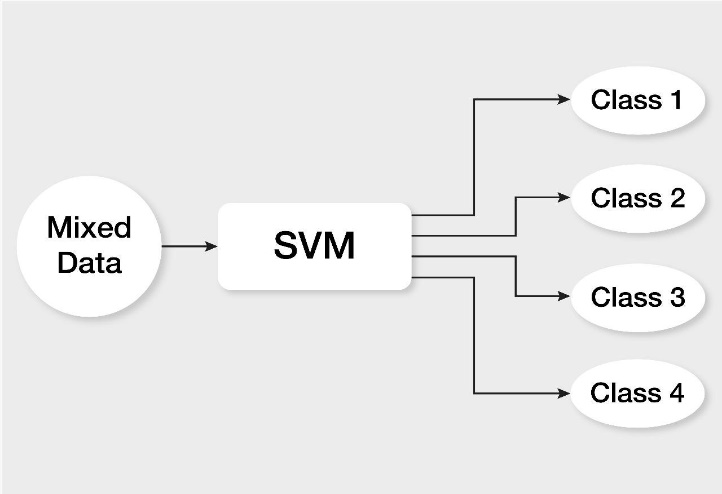
(x2-x1)

The Euclidean distance between (x1, y1) and (x2, y2) is,d = √ ((x2 - x1) ^2 + (y2 - y1) ^2)

So, in short form,

#### 3.5.3 Support Vector Machine

Support Vector Machine (SVM) is an algorithm in machine learning for classification and regression tasks. Its core objective is to find the best possible separation (hyperplane) between different categories or to make predictions based on existing data. It is particularly adept at handling situations where data points belong to other groups and must be divided as clearly as possible. SVMs excel in simple and complex scenarios, offering an excellent tool for various machine learning problems, from image recognition to financial forecasting.

SVM works by finding a hyperplane that maximizes the margin between different classes of data points in a high-dimensional space. It can handle both linear and non-linear classification problems through the use of kernel functions, making it a versatile and powerful machine learning algorithm for various applications.

**FIGURE: 3.2.1 PROPOSED METHODOLOGY OF WO**

FIGURE: 3.5.2.1 Fuzzy Logic

Figure 3.5.3.1 SVM Classes



Figure 3.5.3.2 SVM Visualization

#### 3.5.4 Decision Tree

A decision tree is a versatile machine learning algorithm used for classification and regression tasks. It constructs a tree-like structure where each internal node represents a feature-based decision, and each leaf node provides a predicted outcome. The algorithm recursively partitions the data column into subsets, choosing the best feature at each step to maximize the separation between classes (for classification) or minimize error (for regression). Decision trees are interpretable and capable of capturing complex relationships in the data, but they can be prone to overfitting. Techniques like pruning and setting appropriate parameters help mitigate this issue, and decision trees serve as the foundation for ensemble methods, such as Random Forests, which enhance predictive accuracy by combining multiple decision trees. We have used classification approach.

Decision trees for categorical datasets:

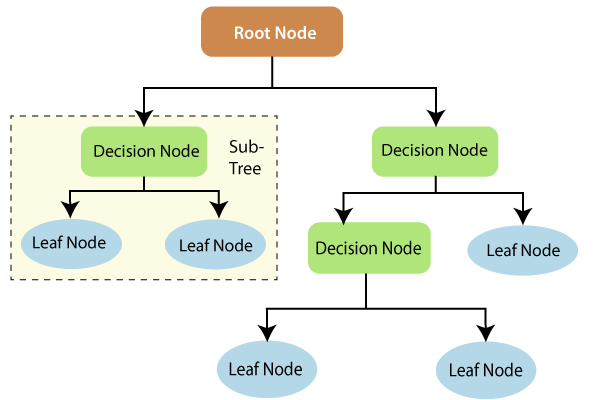
* Choose the best categorical attribute to split the data (e.g., using information gain)
* Split the dataset into subsets based on attribute categories.
* Repeat this process recursively for each subset.
* Traverse the tree to predict new data points based on attribute decisions

Figure 3.5.4.1 Decision Tree

Formula:

Although there are other approaches to choosing the optimal attribute at each node, the Gini impurity and information gain methods are the two most frequently used as a splitting criterion in decision tree models. They aid in assessing the effectiveness of each test condition and its capacity to categorize samples into a group.

Entropy:

Where,

* S stands for the collection of data from which entropy is determined.
* The classes in the set are represented by c,
* S p(c) denotes the percentage of the set's total number of data points that are members of class c, S.

Information Gain:

Information Gain = entropy (parent) – [average entropy (children)]

Gini: The gini impurity is calculated by this formula-

Where, Pj is the probability of class j.

3.6 The background and rationale of design choice

K-Means is a clustering algorithm introduced by Stuart Lloyd in 1957 at Bell Telephone Laboratories. It's designed for grouping data points into clusters based on similarity, with the objective of minimizing the sum of squared distances between data points and their cluster centroids. Regional Origin: K-Means doesn't have a specific regional origin, as it's a mathematical algorithm widely used in data analysis and machine learning across the globe.

SVM was developed by Vladimir Vapnik and Alexey Chervonenkis in the 1960s, with significant advancements in the 1990s. Initially a binary linear classification algorithm, SVM has evolved into a versatile tool for both classification and regression tasks, emphasizing maximizing the margin between classes.

Regional Origin: The founders of SVM, Vladimir Vapnik and Alexey Chervonenkis, have backgrounds from the former Soviet Union (Russian and Ukrainian, respectively). However, SVM's use and influence extend globally in machine learning and data science.Decision trees have roots in statistics and computer science. They were formalized for decision-making by Morgan and Sonquist in 1963 and further developed, particularly by Leo Breiman and others, in the 1980s. Decision trees are used for classification and regression tasks, creating a tree-like structure of decisions to make predictions.

Regional Origin: Decision trees, as a concept, do not have a specific regional origin. They have been developed and improved by researchers and practitioners from various regions and are widely used in machine learning and data analysis globally. The Naive Bayes classifier is based on Bayes' theorem, which was developed by the English statistician Thomas Bayes in the 18th century. The "naive" part of Naive Bayes comes from the assumption of independence between features in the classification process, which simplifies calculations.

Regional Origin: Bayes' theorem, which forms the foundation of Naive Bayes, was developed by Thomas Bayes, an English mathematician. The algorithm itself is not tied to a specific region and is used worldwide for text classification, spam filtering, and more.

3.7 Requirements

Once we carefully examined all the essential statistical and theoretical concepts and techniques, we compiled a roster of the required hardware, software, and development tools necessary for forecasting career recommendations.

Hardware/Software Requirements:

* Operating System (Windows 7 or more)
* Ram (4 GB or more)
* Web Browser ( chrome or any suitable one)

Developing Tools

* Python 3.10 or latest version
* Anaconda
* Visual Studio code
* Jupiter notebook

**Chapter IV. Result and Discussion**

This chapter will cover the details of our research, including the experimental results. We will compare our algorithms to select the most effective one for prediction, and visualize the output using a confusion matrix. To conclude the chapter, we will summarize the results.

4.1 Experimental Results

We utilized more than 10000 simulated datasets from various fields of Computer Science and Engineering and crafted different questions to evaluate the recommended field of study for a student's future career and job. The dataset comprises ten diverse attributes covering several questions related to a student's personal life and interests. The output of this analysis is a recommendation of a student's field of study based on their interests and previous academic pursuits, specific to the Bangladeshi environment. Our output class for the system is a student's chosen discipline or field of study, with ten fields. We employed two machine learning models, namely Naïve Bayes and k-NN, Decision Tree and SVMt o analyze our datasets and generate recommendations. Although the initial results were unsatisfactory, a real-world dataset is expected to enhance the accuracy of the output.

4.2 Prediction of Results

The machine learning model's performance can be assessed using a confusion matrix, which is a table consisting of four values: true positive, true negative, false positive, and false negative. This technique can be applied to the test dataset, where the true values are already known. True positives (TP) are the values where the examples are accurately identified as positive. False positives (FP) are the values where the examples are negative, but they are mistakenly classified as positive. True negatives (TN) are the values where the examples are accurately identified as negative. Finally, false negatives (FN) are the values where the examples are positive, but they are incorrectly classified as negative.

Performance measures can be calculated using the following formulas:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Classification Model | Accuracy | Precision | Recall | F-measure |
| Decision Tree | 96.92 | 99.88 | 100.0 | 100.0 |
| SVM | 92.68 | 99.581 | 99.588 | 99.647 |
| Naïve Bayes | 89.84 | 99.82 | 99.88 | 99.94 |
| K-NN | 90.5 | 96.53 | 96.11 | 96.29 |

TABLE 4.2.1: PERFORMANCE MEASUREMENTS

Based on the table presented above, the Decision Tree Classifier algorithm has been identified as the most effective classifier for providing recommendations regarding suitable choice.

Figure 24.2.1 Accuracy of the Models

**FIGURE: 3.2.1 PROPOSED METHODOLOGY OF WO**

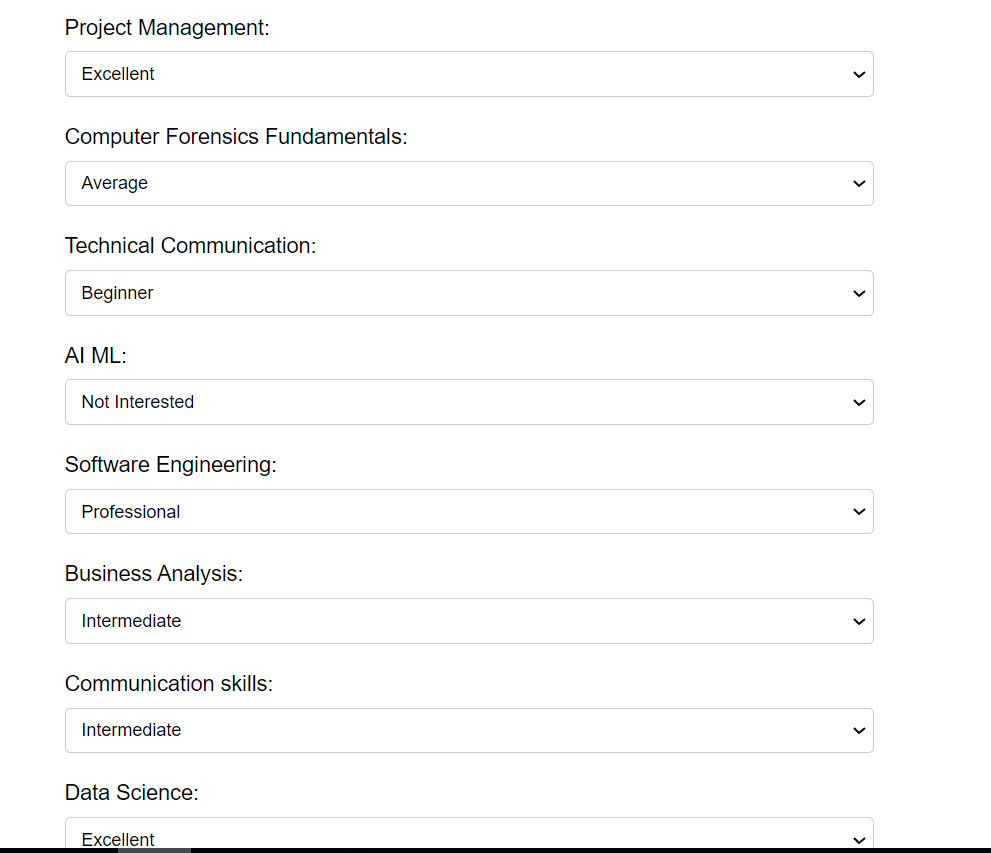
FIGURE: 4.2.1: Accuracy

F=-Measures values are mentioned below:

TABLE 4.2.2: F-MEASURE OF CATEGORY

|  |  |
| --- | --- |
| Categories | F-Measure |
| AI ML Specialist | 98.5 |
| API Specialist | 98.5 |
| Application Support Engineer | 100.0 |
| Business Analyst | 94.5 |
| Customer Service Executive | 95 |
| Cyber Security Specialist | 95 |
| Data Scientist | 96 |
| Database Administrator | 96 |
| Graphics Designer | 98 |
| Hardware Engineer | 99 |
| Helpdesk Engineer | 100 |
| Information Security Specialist | 98 |
| Networking Engineer | 97 |
| Project Manager | 97 |
| Software Developer | 98 |
| Software tester | 98 |
| Technical Writer | 98 |

#### 4.2.1 Input Snapshot:



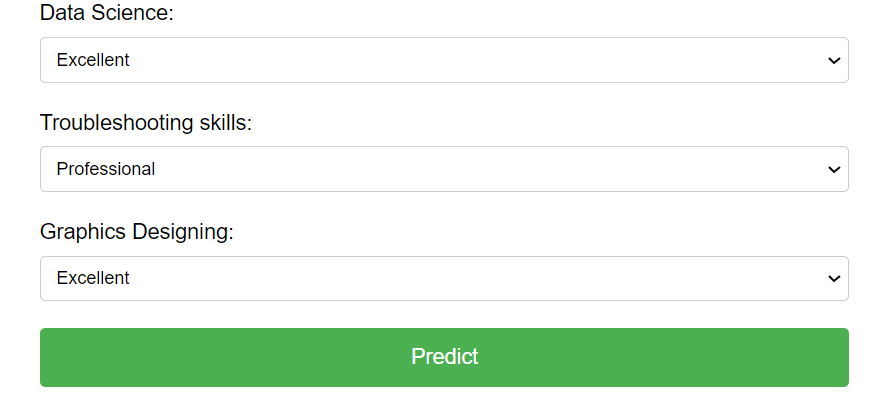


Figure 4.2.1.1 Input

#### 4.2.2 Output Snapshot

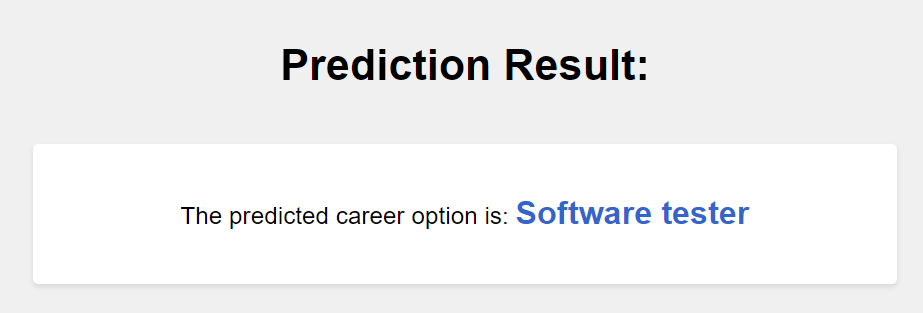


Figure 4.2.1.2 Output

#### 4.3 Significance of Research

The research on career guidance recommendation system holds significant importance in today's rapidly changing and competitive job market. With the increasing number of fields and specializations, it can be overwhelming for students to make the right career choice. A recommendation system based on machine learning algorithms can provide personalized career guidance to students, which can help them make informed decisions about their future. The significance of this research lies in its potential to benefit students, educational institutions, and industries. By providing students with personalized recommendations based on their interests and academic history, the system can assist them in selecting the most suitable career path. Educational institutions can also use this system to enhance their career guidance services and ensure that their students are well-equipped for the job market. Additionally, industries can benefit from a more skilled and specialized workforce, which can ultimately lead to economic growth. Moreover, this research can contribute to the development of machine learning algorithms and techniques, which can be applied to various other fields. It can also provide insights into the factors that influence career choices and help researchers and policymakers develop strategies to address issues related to unemployment and skill gaps. Overall, the research on career guidance recommendation systems can have a significant impact on the education and job sectors, the economy, and the advancement of machine learning techniques. It can ultimately benefit society as a whole by providing individuals with the guidance they need to succeed in their chosen careers.

4.4 Limitations

We are working with artificial datasets but real world datasets are better for prediction. If we use artificial datasets our prediction can be or results can be different. Our research is only for the fields of Computer Science and Engineering Departments. We have not covered other departments like CE, ME, EEE etc. Also we have not covered Medical and Business field.

4.5 Implementation for further study

There are several implications for further study that can be from our research:

* Data sets can cover the real world data sets and departments can be other engineering departments, Business and Medical fields.
* Since our research was conducted using dummy datasets, one important future direction would be to test our approach on real-world datasets to evaluate its effectiveness in practical settings.
* We can also consider adding more fields of study to our research work to provide a wider range of recommendations to students.
* Another potential avenue for further study is to deploy our research as an Android and website application to make it more accessible and user-friendly for students seeking career guidance.
* Lastly, we can explore the possibility of publishing a publication that details our research work to disseminate our findings to a wider audience and invite further discussion and collaboration in the field.

**Chapter V. Conclusion**

5.1 Conclusion

The career guidance recommendation system is designed for Bangladeshi students who struggle with choosing a subject for their graduation. By using machine learning models like Naïve Bayes, k-NN, and Fuzzy Logic, the system recommends a field of study based on a student's interests and previous academic pursuits. This research can assist students in selecting a field of study and can benefit industries that employ these students in the future. Furthermore, the system has the potential to be extended to other fields of study and deployed as a mobile application or website. This can greatly benefit students in making informed decisions about their career paths and help them achieve their desired career goals. The future work for this research includes using real-world datasets to further validate the accuracy of the system, adding more fields of study, and improving the system's user interface. Overall, the Career Guidance Recommendation System using Machine Learning can have a significant impact on improving the career guidance services provided to students, and help them make informed decisions about their future career paths.

**References**

* Vignesh, S., Priyanka, C.S., Manju, H.S. and Mythili, K., 2021, March. An intelligent career guidance system using machine learning. In 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 987-990). IEEE.
* Roy, K.S., Roopkanth, K., Teja, V.U., Bhavana, V. and Priyanka, J., 2018. Student career prediction using advanced machine learning techniques. International Journal of Engineering & Technology, 7(2.20), pp.26-29. [3] Metha S., Kamble S., Kumar. (2021) 'Career Guidance using Machine Learning.' 9(VI) pp. 4124-4127.
* Metha S., Kamble S., Kumar. (2021) 'Career Guidance using Machine Learning.' 9(VI) pp. 4124-4127.
* Ajay, D., Nilesh, S., Anirudh, D., Huzaif, S. and Nagare, P.V.(2022) 'AN INTELLIGENT CAREER GUIDANCE SYSTEM USING MACHINE LEARNING.' 9(3) pp. 1684-1691.
* sayyed F.,Sanghani R., Vora A., Lemos. (2020) 'AI based Career Guidance.' 7(5) pp. 2505-2509
* Copeland, L.Y., Dik, B.J., McLaren, M.R., Onder, C., Wolfson, N.E. and Kraiger, K., 2011. Recommendations for using computer‐assisted career guidance systems (CACGS) in career counseling practice. *Journal of Psychological Issues in Organizational Culture*, *2*(3), pp.86-94.
* Gati, I., Levin, N. and Landman-Tal, S., 2019. Decision-making models and career guidance. *International handbook of career guidance*, pp.115-145.
* Roy, K.S., Roopkanth, K., Teja, V.U., Bhavana, V. and Priyanka, J., 2018. Student career prediction using advanced machine learning techniques. International Journal of Engineering & Technology, 7(2.20), pp.26-29. [3] Metha S., Kamble S., Kumar. (2021) 'Career Guidance using Machine Learning.' 9(VI) pp. 4124-4127.
* Nabil Ahmad, S., 2020. Personalized Career Path Recommender System for Engineering Student. *IEEE,* 8 October , 6(14), p. 9.
* Vignesh, A. H. N. H., 2021. Career Path Recomendation System. *IEEE,* 17 June, 13(16), p. 7.