LONDON METROPOLITAN UNIVERSITY

islington college
(इस्लिङ्टन कलेज)

Module Code & Module Title

Level 6 – Artificial Intelligence

Assessment Type 25%

Proposal

Semester

2024/25 Autumn

Student Name: Sumit Pokhrel

London Met ID: 22085636

College ID: NP01CP4S230034

Assignment Due Date: Wednesday, December 25, 2024

Assignment Submission Date: Wednesday, December 25, 2024

Submitted To: Dipeshor Silwal

Word Count (Where Required):

I confirm that I understand my coursework needs to be submitted online via MST Classroom under the relevant module page before the deadline in order for my assignment to be accepted and marked. I am fully aware that late submissions will be treated as non-submission and a mark of zero will be awarded

22085636 Sumit Pokhrel.docx.docx

Islington College,Nepal

Document Details

Submission ID
trn:oid:::3618:74747702

Submission Date
Dec 25, 2024, 1:16 AM GMT+5:45

Download Date
Dec 25, 2024, 1:18 AM GMT+5:45

File Name
22085636 Sumit Pokhrel.docx.docx

File Size
22.6 KB

25 Pages

3,445 Words

19,077 Characters

Match Groups

- **31** Not Cited or Quoted 12%
  Matches with neither in-text citation nor quotation marks
- **14** Missing Quotations 4%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **1** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

6% Internet sources
3% Publications
15% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1 Submitted works
Liverpool John Moores University on 2024-09-09 3%

2 Submitted works
Southern Cross University on 2023-03-12 1%

3 Submitted works
Wilmington University on 2024-11-18 1%

17% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Match Groups

- **31** Not Cited or Quoted 12%
  Matches with neither in-text citation nor quotation marks
- **14** Missing Quotations 4%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **1** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

6% Internet sources
3% Publications
15% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Table of Contents

1.Introduction 1
1.1 Introduction to Arttifical Intelligence (AI) 1
1.1.1 How does AI works 1
1.2 Machine Learning 1
1.2.1Types of Machine Learning 2
1.3 Introduction to the Income Prediction system 3
1.4 Explanation of the AI algorithm concepts used 3
1.4.1Research work done method 3
2.Background 5
2.1 Similar system case study 5
2.2 Dataset 8
3.Solutions: 10
3.1 Approach to solving the problem 10
3.2 Explanation of the algorithm used 10
3.3 Pseudocode 15
3.4 Flowchart 16
4.Conclusion 17
4.1 Analysis of work done 17
4.2 How does the solution address real problems? 17
4.3 Future work 18
5.References 19

Table of Figures

Figure 1:Supervised Learning algorithms 3
Figure 2: Income Prediction Using Machine Learning Techniques 5
Figure 3:Adult Income Analysis 6
Figure 4:Predicting Adult Census Income with Machine Learning Techniques 7
Figure 5:Mathematical explanation of Logistic regression 11
Figure 6:Gini Impurity 12
Figure 7:Entropy 12
Figure 8:Information Gain (IG) 12
Figure 9:Euclidean distance 14
Figure 10: Manhattan Distance 14
Figure 11:Flowchart 16

Table of Tables

Table 1:Dataset.. 8

1.Introduction

1.1 Introduction to Arttifical Intelligence (AI)

Artificial intelligence (AI) can be defined as the ability of the computer or a computer operated device to make decisions as those can be made by intelligent entities. This term is often applied to the challenge of developing systems that have the capabilities that set human beings apart — the abilities to reason, to make meaning, to generalize, and to learn from experience. Clerks On the other hand digital computers since the 1940s have been-built to do extremely complex things like playing World Chess or finding proofs for the theorem. Despite the fact that computer performance and storage capability are growing rapidly, there is as yet no program that can retain as wide a flexibility as man in topics related to a larger amount of information as could be useful for such knowledge-oriented activities. (geeksforgeeks, 2024)

1.1.1 How does AI works

Branch of computer science which is also known as artificial intelligence, or simply as AI, is the scientific field focusing on development of intelligent systems able to learn from experience as well as execute specific tasks without further coding. It was only possible for people to do the things, which is today's achievement of AI is that computers can do all these things. AI develops a number of approaches such as robotics, machine learning, neural networks and natural language processing to fulfill its goals. (geeksforgeeks, 2024)

1.2 Machine Learning

A subfield of artificial intelligence known as machine learning enables computers to identify concealed patterns within sets of data so that they can assess results on new similar data without being programmed for every task individually. Traditional statistical learning tasks even produce valuable knowledge by using mathematical models to make predictions. Digital services based on this approach can be selected for the following fields: recommendations, protection against fraudulent operations, portfolio, natural language processing, picture and audio recognition, and job automation. (geeksforgeeks, 2024)

1.2.1Types of Machine Learning

a) Unsupervised Machine Learning

There is the kind of machine learning called supervised learning, in which the labeled dataset is employed to train the model. The nature of topology design involves it uses labeled training data to learn the relation between input characteristics and targets. In supervised learning, it evolves the ability to predict from the form of characteristics and the corresponding labels of any other data it has not viewed before. (geeksforgeeks, 2024) There are two types of Unsupervised Machine Learning are as follow:

- **Clustering:** There is one kind of machine learning that used to arrange the large amount of unlabeled dataset namely clustering or cluster analysis. By definition, it is “a technique of placing the data points into different clusters where each cluster consists of similar data points. The items as to which one could be related, remain in a group that is either less related or is not similar to another group. (geeksforgeeks, 2024)
- **Dimensionality reduction:** Due to the fact that most models aim at reducing the amount of information loss, dimensionality reduction algorithms seek to reduce the number of input variables in a given set. One way of making real datasets easier to visualize and analyze is by decreasing its complexity. The well-known methods of this class are Principal Component Analysis (PCA), t-Distributed Stochastic Neighbor Embedding (t-SNE), and autoencoders. (geeksforgeeks, 2024)

b) Supervised machine learning:

Supervised machine learning uses prepared data with correct training data that are fed to the machines, to enable the machines to make predictions. As noted by the labelled data, some of the input data has been prelabeled in terms of the output. (javatpoint, 2024) Types of Supervised Machine Learning are as follow:

- **Regression:** In cases where the input and output variables are related, the system uses the regression processes. They are applied in the prediction of changeable scales such as in the market trend and future weather conditions. (geeksforgeeks., 2024)

- **Classification:** Classification algorithms are used when the target variable is qualitative and when the value of the dependent variable can take only two values: Yes/No, Male/Female, True/False, etc. (geeksforgeeks, 2024)

1.3 Introduction to the Income Prediction system

Income prediction system is a conceptual framework for user to make a decision on income level of a particular person. The use of income level prediction system is very quiet useful, for Individual user Income has negative reaction in the world it is related with the social and political crisis that leads to mental illness and low life expectancy. To rectify this problem the income prediction system works a foundation in every field including the politics and economics. The Income Prediction system is a solution to the income problem. Income estimate is therefore useful in the prediction of risk-prone population and allocation of resources appropriately.

1.4 Explanation of the AI algorithm concepts used

1.4.1Research work done method

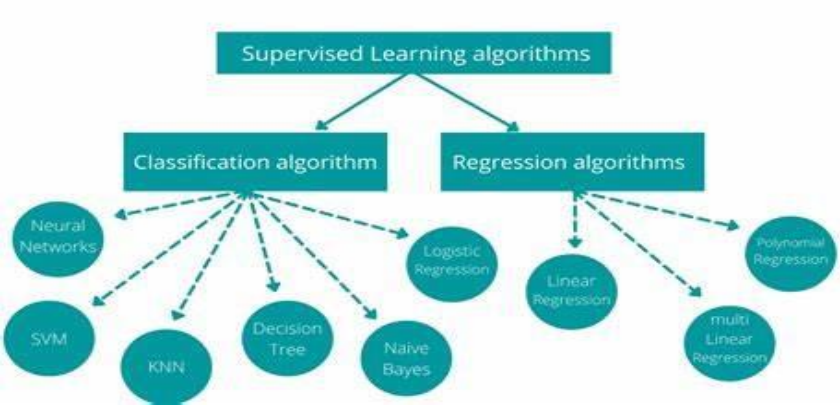


Figure 1:Supervised Learning algorithms

Classification:

Classification in a form of supervised learning approach involves an algorithm being trained on a labelled training set with the view of classifying fresh, unlabelled data. The main objective of classification machine learning is to develop a method that will categorize or assign a new observation into a right class based on its features. (geeksforgeeks, 2024)

- **K-Nearest Neighbors**: KNN is one of the simplest yet most important classification algorithms in the general spectrum of machine learning algorithms. This algorithm is classified under the supervised learning, and it is applied mainly in intrusion detection, data mining and pattern recognition. (geeksforgeeks, 2024)

- **Decision Trees**: A decision tree is a kind of diagram that is being used in the cases of forecasting or in the decision-making process. Leaf nodes are and when you have the final result or forecast, Branches portray the result of these decision and Nodes portray decision or attribute test. (geeksforgeeks., 2024)

- **Logistic Regression:** Logistic regression is a type of supervised learning algorithm, which used with classification problem in order to predict the probability that an instance belong to a specific class or not. Logistic regression is a statistical tool that in particularly examines the connection of two variables in a data set. The article provides the basics of the logistic regression, its' types and the ways it can be used. (geeksforgeeks, 2024)

2.Background

2.1 Similar system case study

- **Income Prediction Using Machine Learning Techniques**

UNIVERSITY OF CALIFORNIA
Los Angeles

Income Prediction Using Machine Learning Techniques

A thesis submitted in partial satisfaction
of the requirements for the degree
Master of Applied Statistics & Data Science

by

Kahyun Jo

Figure 2: Income Prediction Using Machine Learning Techniques

This system uses a variety of demographic predictor variables, including capital gain, education level, relationship, occupation, and capital loss, this system explores the use of advanced machine learning techniques to predict income levels, specifically whether individuals earn over $50,000 annually. Income level forecasting is essential for explaining economic inequality and guiding policy choices. This study includes a detailed analysis of each model's performance using the Adult Income dataset from the UCI Machine Learning Repository, which includes socio-economic and demographic factors. Following a preprocessing step to guarantee data quality, a variety of machine learning techniques are applied i.e. Support Vector Machines, Neural Networks, Decision Trees, Random Forests, k-Nearest Neighbors, and Logistic Regression. In this system Baseline Forest achieve accuracy 86.410%, specify 88.600%, and RMSE 0.315. Tuned Random Forest achieve AUC 94.964% and F1 score 82.057%. (Jo, n.d.)

- **Adult Income Analysis**

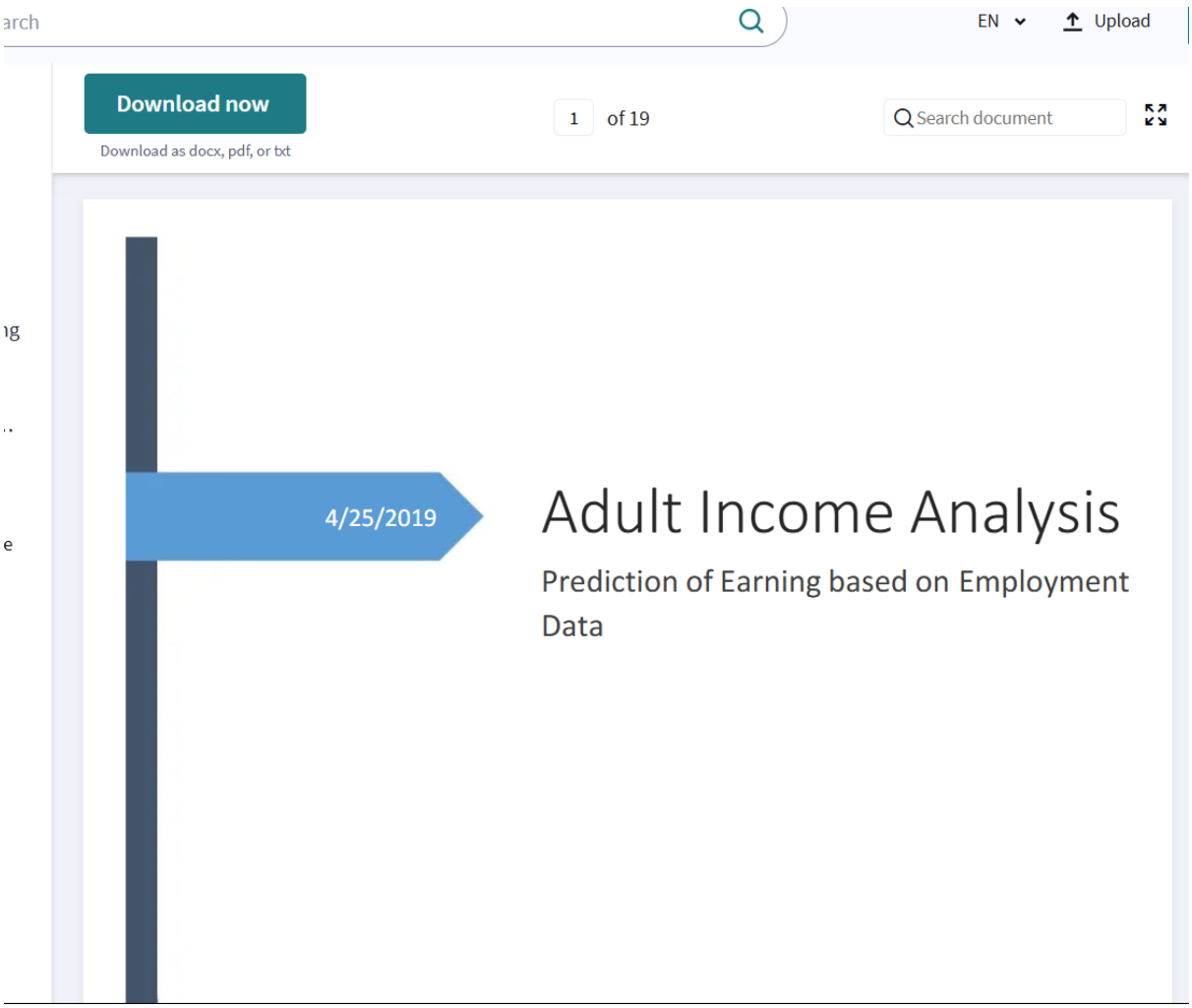


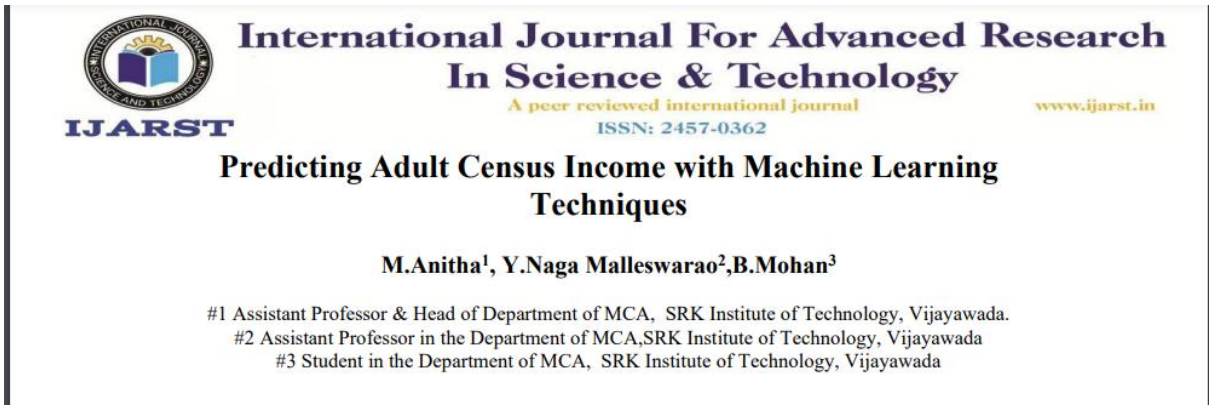
Figure 3:Adult Income Analysis

This system aims to show the usage of machine learning and datamining techniques in providing a solution to income equality problem. In this system Classification has been chosen for predict. In the classification has to predict whether the person income is more than 50,000or less than50,000 us dollar. System provides different types of classification i.e.

- ✓ Navies Bayes algorithm
- ✓ J48 Decision Tree
- ✓ Logistic Regression

From the uses of the algorithm Logistic Regression has the highest accuracy of 93% and Navie Bayes algorithm Give the Accuracy of 83%. (PATEL, 2022)

- **Predicting Adult Census Income with Machine Learning Techniques**



International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

Predicting Adult Census Income with Machine Learning Techniques

M.Anitha[1], Y.Naga Malleswarao[2],B.Mohan[3]

#1 Assistant Professor & Head of Department of MCA, SRK Institute of Technology, Vijayawada.
#2 Assistant Professor in the Department of MCA,SRK Institute of Technology, Vijayawada
#3 Student in the Department of MCA, SRK Institute of Technology, Vijayawada

Figure 4:Predicting Adult Census Income with Machine Learning Techniques

The application of Machine Learning in financial as well as the related data analysis is gradually growing in use. Applying the concept of ML has now become almost compulsory in most industries for ascertaining the probabilities of the likely revenues. This system is concerned with the financial sector, that is, categorizing people that earn over 50, 000US dollar per year. Such binary classification problems are illuminating and rather essential in a range of fields, such as accounting, online marketing, and even socioeconomic evaluation estimated budget. The Adult Census Dataset, which was imported from Kaggle Dataset, is the main data used in this system. Universally the dataset is very helpful for in depth and modeling-based research because of the qualitative nature of the type of data that is featured and these are income, employment, education and demographic of the people. The research establishes the best option depending on the evaluation of the performance of the built models via suitable criteria thorough evaluation procedures. Whenever a precise model of income levels needs to be developed for the given input data, then the refining and optimization methods are applied. This system used four types of algorithms of classification i.e.

- ✓ Decision Tree: accuracy score of the model is 0.8061
- ✓ Random Forest: accuracy of the model is 0.8504
- ✓ XG boost: accuracy of the model is 0.7582
- ✓ KNN: accuracy of the model is 0.7786

So, from the accuracy of the algorithm used in this system Random Forest is the best algorithm in the given problem. (M.Anitha1, n.d.)

2.2 Dataset

A dataset is helpful information that developers probably need for the achievement of their goals. The number of records present in a dataset is its rows, while the features present in a dataset make its columns. With the help of them, they are mainly used to teach machines, gather information, and make reasonable conclusions as to the corresponding fields such as government, industry, machine learning, etc. Data size and scope may differ, and in many cases, to ensure data credibility and suitability to analysis and modeling, data needs to be cleansed and transformed in one manner or another. (geeksforgeeks, 2024) Income prediction system dataset is given on the Kaggle website(Datasetlink)

It is a website that provides datasets for numerous things users to produce their datasets. The table of dataset are as follows:

Table 1:Dataset

Dataset	Data Type	Description
Age	Float64	The age of individual.
Work class	object	The type of work the individual is engaged in, such as Private, Self-Employed, or Government work.
Fnlwgt	Float64	Final weight (an indicator used in survey data to adjust for population estimates).
Education	object	The highest level of education attained (e.g., Bachelors, HS-grad, Doctorate).
Education-num	Float64	A numeric representation of the highest education level attained.

Martial-status	object	The marital status of the individual (e.g., Married, Divorced, Single).
Occupation	object	The type of occupation the individual is engaged in (e.g., Sales, Craft-repair, Tech-support).
Relationship	object	The relationship status of the individual within a household (e.g., Husband, Wife, Unmarried).
Race	object	The race of the individual (e.g., White, Black, Asian-Pac-Islander).
sex	object	The gender of the individual (Male or Female).
Capital Gain	Float64	Capital gains income earned by the individual.
Capital Loss	Float64	Capital losses experienced by the individual.
Hours per Hour	Float64	The number of hours the individual works per week (continuous).
Native Country	object	The country of origin of the individual (e.g., United-States, Mexico, Canada).
Income	object	Binary income category: <=50K or >50K per year.

3.Solutions:

3.1 Approach to solving the problem

Although the prediction system has been widely researched.it is still necessary to make income prediction easier with the growing importance of income classification for applications such as market segmentation, financial services, and personalized recommendations, it becomes essential to develop effective and reliable income prediction systems There are challenges in creating accurate and efficient income prediction systems: Here are some solutions to these problems:

- **Handling Missing Data**: Missing data can impact on system performance .data cleaning is necessary for the system.so missing data can be remove from dropna function.
- **Improve accuracy**: Decision Trees and Random Forest are particularly effective for capturing non-linear relationships in the data. Random Forest is very suitable Algorithm For improve Accuracy of the system.
- **Scable**: KNN algorithm although works very well if the data set is small in size is very expensive if the data size is large. Decision Trees and Random Forest are better scaling techniques and more appropriate for large datasets making them even more practical for real-world scenarios.

3.2 Explanation of the algorithm used

In the Income prediction system, the algorithm used are as follows

i. **Logistic regression:**

A specific type of classification technique known as logistic regression is applied on the given classification problems if the aim is to estimate the probability that an instance belongs to a particular class or not. Logistic regression is a procedure that discriminative the relation of two variables. The article looks at the types, use, and principles of logistic regression. (geeksforgeeks, 2024)

- Key Points for used Logistic regression in project
    a. Logistic regression gives either Yes or No, true or False.it predicts the probability that input belongs ta a specific class so, it predict <=50Kor>50K income in system. (geeksforgeeks, 2024)

b. Logistic regression is used in the prediction of the output of a categorical dependent variable. Therefore, the result needs to be a discrete or category value. (geeksforgeeks, 2024)

- Mathematical explanation

  Logistic regression uses the logistic function (sigmoid Function) to map the predicted values to probabilities.

$$P(y = 1|x) = \sigma(z) = \frac{1}{1 + e^{-z}}$$

where $z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n$, and $\sigma(z)$ is the sigmoid function.

β_0 is the intercept, and β_i are the coefficients for the i-th feature x_i.

Figure 5:Mathematical explanation of Logistic regression

Therefore, in income Prediction system there is only <=50Kor>50K income to predict it same like0or1 or YesorNo.so, Logistic regression is suitable for income prediction system.

ii. Decision tree:

A decision tree is a modeled graphical tree that offers prediction or decision-making and compartments closely resemble flowcharts. It consists of decision and/or attribute test nodes, branches, which represent how it has gone, and end nodes which have the result or the prediction. Any internal node corresponds to an attribute test; any branch corresponds to the test result; any leaf node corresponds to a class value or a continuous value. (geeksforgeeks., 2024)

- Key Points for used Logistic regression in project
  a. The process of using decision trees is easy and the results are easy to analyze. The manner in which it is illustrated actually resembles how people make decisions.
  b. Decision trees do not require normalization or scaling of the data.

- **Mathematical explanation**
  - **Gini Impurity:** estimates a random guess for the new instance through computing the quantity that the new instance would be put in the wrong category if it as well as the ratio of classes in the training data set are employed to guess randomly. (geeksforgeeks., 2024)

$$G = 1 - \sum_{i=1}^{C} p_i^2$$

where p_i is the proportion of samples belonging to class i.

Figure 6:Gini Impurity

- **Entropy**: Measures the amount of uncertainty or impurity in the dataset.

$$H = -\sum_{i=1}^{C} p_i \log_2(p_i)$$

Figure 7:Entropy

- **Information Gain (IG)**: Measures the reduction in entropy or Gini impurity after a dataset is split on an attribute**.**

$$IG = H_{\text{parent}} - \sum_{\text{children}} \frac{N_{\text{child}}}{N_{\text{parent}}} H_{\text{child}}$$

Figure 8:Information Gain (IG)

iii. K-Nearest Neighbors (KNN)

KNN is one of the simplest but among the most important machine learning classification algorithms. This technique comes under the type of Supervised learning and is extensively used in Intrusion detection, Data mining and Pattern recognition. In contrast to GMM, for example, this approach does not presuppose the Gaussian distribution of the provided data; this method is broadly used in actual-life situations because the approach does not make any assumptions about the nature of the data distribution. Sometimes, we are given some sample data which is called training data to unravel how coordinates are arranged depending on specific characteristics. (geeksforgeeks, 2024)

- Key Points for used Logistic regression in project
  - K-NN algorithm is one of the more popular and flexible algorithms in the class of machine learning methods and is used mainly for its simplicity and ease of application. Concerning the distribution of the underlying data, no assumptions are made in this method. Moreover, it also works with categorical variables as well as numerical, making it a flexible choice regarding a kind of dataset in tasks including regression and classification. This procedure produces predictions depending on the extent to which data points in a given dataset are similar to one another; it is a non-parametric one. Compared to other algorithms K-NN is not very sensitive to outliers. (geeksforgeeks, 2024)
  - This method, called K-NN, means that you use distance such as Euclidean distance and identifies the K nearest points to a certain point. The resulting value is calculated using the average of the K neighbors or the simple majority by assigning class or value to the data item. In this way, the algorithm may reconfigure to different structures and provide predictions based on the local map of the data. (geeksforgeeks, 2024)

- **Mathematical explanation**

  As is well known, the KNN algorithm helps to find groups or the nearest neighbors of a given query point. Nonetheless, we need a measure in order to find the nearest points or clusters for a query point. We employ the distance metric listed below for this purpose:

- **Euclidean distance**:

  This is simply the cartesian distance that can exist between the two generalized locations in the plane or hyperplane here. The distance, which can be measured to represent the Euclidean distance is also represented by the distance of the straight-line passing through the two sites under consideration. This measure maybe used to evaluate the net displacement between two states of an object. (geeksforgeeks, 2024)

$$\text{distance}(x, X_i) = \sqrt{\sum_{j=1}^{d} (x_j - X_{i_j})^2]}$$

Figure 9:Euclidean distance

- **Manhattan Distance:**

  This is mostly used when we are interested at the total distance of the move rather than displacement of the object. This medium is arrived at by simply summing the absolute value of the differences of the points' coordinates that are in n dimensions. (geeksforgeeks, 2024)

$$d(x, x') = \sum_{i=1}^{n} |x_i - x_i'|$$

Figure 10: Manhattan Distance

3.3 Pseudocode

In algorithm development, pseudocode is central to the understanding of the problem approach by the reader and the programmer and preparation of the solution. A recent intermediate form of an algorithm and a program, pseudocode helps in conversion of the algorithm into the program. (geeksforgeeks, 2024)

START
IMPORT the libraries
Load dataset from a CSV file
Check for missing values
Handle missing values
Remove duplicate values from the dataset
Split the dataset into training and testing sets
Scale numerical features
TRAIN the model logistic regression
Evaluate the datasets
END

3.4 Flowchart

The flowcharts are just to represent the data or algorithm that can help the user to understand the program. It demonstrates procedure for solving an algorithm, method or a problem practically on how it ought to be done. Contrary to the to-do lists style, it is a graphical way to display steps that most beginners prefer in comprehending Computer science algorithms which assist in fixing algorithmic issues. A flowchart is a diagram of boxes which represents a pattern of operation by showing a Record of work flow. (visual-paradigm, 2024)

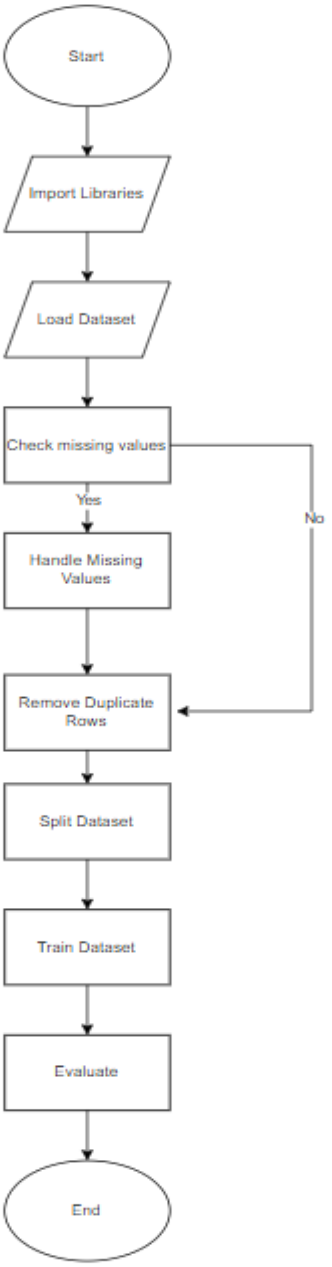


Figure 11:Flowchart

4.Conclusion

4.1 Analysis of work done

An income prediction system is am mechanism for shifting data from big datasets and predict income based on their data. Every day, there is tremendous growth in the amount of data in of individual income, making it difficult for company, people or so on to predict their individual income or company employee income. Income prediction system predict income depending on the data supplied by the person or company. Logistic regression, Decision Tree or KNN are simple, scable and reliable.

As for the income prediction system it is referred to go with the KNN algorithm, Decision Tree, and Logistic Regression. The peculiarities of the managerial application of these techniques, enhancing the handling of different data complexity, enable the usage of these techniques for working with revenue prediction problems.to boss the system's scability and scarcity, advanced computational methods are necessary. This solution addresses this algorithm's deficiencies as well as enhance the performance of the system. To generate improved all recommendations, many tactics are used, for example, reducing an error rate and raising efficiency.
accuracy and precision.

4.2 How does the solution address real problems?

prediction system is utilized in wide range application and have shown to be extremely useful in recent years. These models have become so famous that they are being used in many companies or government office to predict income, salary and so on. Many companies are utilizing the prediction system it predicts income or salary of person to satisfaction. To improve decision making, efficiency, and user satisfaction, a great number of organizations are now applying income prediction systems. These systems are important for many fields including financial, marketing and employment fields. Thus, based on whether the incomes belongs to certain categories or not, companies can make necessary service and products to provide for customers.

4.3 Future work

The developmental process phase, as well as the research phase of this project, has been completed. Here is what I've done in build income prediction system and. Also, the flowcharts, algorithms, and Pseudocodes needed for the project were also designed. Hence, different training and testing with more data and coming up with better algorithms for the right score could be the important term. Either way, it is better to adhere to the same module work in order to opt for the most effective way.

Thus, it would be valuable to embed it for the GUI and the effectiveness of how it will be in future regarding the real-life application.

5.References

apm, 2025. *What is a Gantt chart?.* [Online] Available at: https://www.apm.org.uk/resources/find-a-resource/gantt-chart/

geeksforgeeks., 2024. Decision Tree. *geeksforgeeks.*

geeksforgeeks., 2024. Regression in machine learning. *geeksforgeeks..*

geeksforgeeks, 2024. Getting started with Classification. *geeksforgeeks.*

geeksforgeeks, 2024. Introduction to Machine Learning: What Is and Its Applications. *geeksforgeeks,* 11 Dec.

geeksforgeeks, 2024. K-Nearest Neighbor(KNN) Algorithm. *geeksforgeeks.*

geeksforgeeks, 2024. Logistic Regression in Machine Learning.

geeksforgeeks, 2024. *Software Development Life Cycle (SDLC).* [Online] Available at: https://www.geeksforgeeks.org/software-development-life-cycle-sdlc/

geeksforgeeks, 2024. Supervised and Unsupervised learning. *geeksforgeeks.*

geeksforgeeks, 2024. What is a Dataset: Types, Features, and Examples. *geeksforgeeks.*

geeksforgeeks, 2024. What is Artificial Intelligence?. *geeksforgeeks,* 21 Aug.

geeksforgeeks, 2024. What is PseudoCode: A Complete Tutorial. *geeksforgeeks.*

geeksforgeeks, 2024. *What is RUP(Rational Unified Process) and its Phases?.* [Online] Available at: https://www.geeksforgeeks.org/rup-and-its-phases/

geeksforgeeks, 2024. What is Unsupervised Learning?. *geeksforgeeks.*

javatpoint, 2024. *Supervised Machine Learning.* [Online] Available at: https://www.javatpoint.com/supervised-machine-learning [Accessed 24 Dec 2024].

Jo, K., n.d. *Income Prediction Using Machine Learning Techniques,* Los Angeles: s.n.

M.Anitha1, n.d. *Predicting Adult Census Income with Machine Learning,* s.l.: s.n.

PATEL, M., 2022. *Adult Income Prediction,* s.l.: s.n.

Shammah, N., 2023. What Is Technology? Meaning, Types And Impacts Explained. *icoverage.io,* 23 July.

visual-paradigm, 2024. *Flowchart Tutorial (with Symbols, Guide and Examples).* [Online] Available at: https://www.visual-paradigm.com/tutorials/flowchart-tutorial/ [Accessed 24 Dec 2024].