Credit card Fraud detection

Satendra Lodhi,Sagar(M.P.)

Aim and Objectives

Credit card fraud is a growing problem in the financial industry, with the potential to cause significantfinancial losses to both customers and financial institutions. As a result, there has been a significantamount of research in recent years on developing effective fraud detection systems. These systems rely ona combination of statistical techniques, machine learning algorithms, and deep learning models to identifyfraudulent transactions.One of the most commonly used approaches for credit card fraud detection isrule-based systems. These systems use predefined rules to identify transactions that are deemed suspicious.However, rule-based systems have limitations, as they are only as good as the rules that have beenpredefined, and they may not be able to detect new types of fraud. To overcome these limitations, machinelearning algorithms and statistical techniques have been applied to credit card fraud detection. Thesetechniques are based on analysing transaction-related data, such as the transaction amount, location, andtime, as well as other relevant factors, such as the customer’s transaction history and account details. Inrecent years, deep learning models, such as convolutional neural networks (CNNs) and recurrent neuralnetworks (RNNs), have also been applied to credit card fraud detection. These models have shownpromising results in identifying fraudulent transactions by learning patterns in the data and improving theaccuracy of fraud detection. Overall, credit card fraud detection is a critical area of research in the financialindustry, with significant potential for improving fraud detection rates and reducing financial losses.

Keywords

Credit Card Fraud Detection, Machine Learning Algorithms, Deep Learning Models

Literature Review

1.1 “Credit Card Fraud Detection System”

The paper discusses the problem of credit card fraud and highlights the importance of detecting fraudulent

transactions promptly. The authors then present their proposed model, which involves pre-processing the

transaction data and then applying various machine learning algorithms, such as logistic regression,

decision trees, and random forests, to classify transactions as either fraudulent or non-fraudulent.

The authors evaluate the performance of their model using a dataset of credit card transactions and

compare it to other models, such as neural networks and support vector machines. The results show that

the proposed model performs better than the other models in terms of accuracy, precision, and recall.

The paper also discusses the limitations of the proposed model, such as the need for a large amount of data

and the challenges associated with handling imbalanced datasets. The authors conclude by highlighting the

potential of machine learning techniques in credit card fraud detection and the need for further research in

this area.

1.2 “Credit Card Fraud Detection using Machine Learning Algorithms”

The authors present their proposed model, which involves pre-processing the transaction data and then

applying a CNN for feature extraction and classification.

To improve the performance of the model, the authors also use feature selection techniques to identify the

most relevant features for fraud detection. They evaluate the performance of their model using a dataset of

credit card transactions and compare it to other models, such as logistic regression and decision trees.

The results show that the proposed model performs better than the other models in terms of accuracy,

precision, and recall. The authors also analyse the contribution of different features to fraud detection and

discuss the limitations of the proposed model, such as the need for a large amount of data and the

challenges associated with handling imbalanced datasets.

1.3 “A Research Paper on Credit Card Fraud Detection”

The proposed model involves pre-processing the credit card transaction data and then apply- ing various

machine learning algorithms, such as Decision Trees, Random Forest, K-Nearest Neighbour, Naive Bayes,

and Artificial Neural Networks, to classify transactions as either fraudulent or non-fraudulent.

The authors evaluate the performance of their model using a dataset of credit card transactions and

compare it with other existing models, such as Logistic Regression, Support Vector Machine, and Gradient

Boosting Machine. The results show that the proposed model outperforms the other models in terms of

accuracy, precision, recall, and F1-score.

The paper also discusses the challenges associated with credit card fraud detection, such as the need for

real-time detection, the challenges of handling imbalanced datasets, and the importance of feature selection

for improving the performance of the model.

1.4 “Credit Card Fraud Detection Predictive Modelling”

The paper covers different machine learning techniques such as supervised, unsupervised, semi-supervised,

and deep learning, and how they are applied in credit card fraud detection. The authors provide a detailed

explanation of each technique, including its advantages and limitations, and also present a comparative

analysis of various machine learning techniques in terms of their performance metrics.

The paper also discusses the challenges associated with credit card fraud detection, such as imbalanced

datasets, the need for real-time detection, and the importance of feature selection for improving the

performance of the model.

The authors conclude the paper by highlighting the potential of machine learning techniques in credit card

fraud detection and the need for further research in this area.

1.5 “A machine learning based credit card fraud detection using the GA algorithm

for feature selection”

The paper covers different deep learning techniques, such as deep neural networks, convolutional neural

networks, recurrent neural networks, and auto encoders, and how they are applied in credit card fraud

detection. The authors provide a detailed explanation of each technique, including their advantages and

limitations.

The paper also presents a comprehensive analysis of different studies that have used deep learning

approaches for credit card fraud detection. The authors provide a detailed summary of each study,

including the dataset used, the deep learning technique applied, and the performance metrics obtained

1.1 “Credit Card Fraud Detection System”

The paper discusses the problem of credit card fraud and highlights the importance of detecting fraudulent

transactions promptly. The authors then present their proposed model, which involves pre-processing the

transaction data and then applying various machine learning algorithms, such as logistic regression,

decision trees, and random forests, to classify transactions as either fraudulent or non-fraudulent.

The authors evaluate the performance of their model using a dataset of credit card transactions and

compare it to other models, such as neural networks and support vector machines. The results show that

the proposed model performs better than the other models in terms of accuracy, precision, and recall.

The paper also discusses the limitations of the proposed model, such as the need for a large amount of data

and the challenges associated with handling imbalanced datasets. The authors conclude by highlighting the

potential of machine learning techniques in credit card fraud detection and the need for further research in

this area.

1.2 “Credit Card Fraud Detection using Machine Learning Algorithms”

The authors present their proposed model, which involves pre-processing the transaction data and then

applying a CNN for feature extraction and classification.

To improve the performance of the model, the authors also use feature selection techniques to identify the

most relevant features for fraud detection. They evaluate the performance of their model using a dataset of

credit card transactions and compare it to other models, such as logistic regression and decision trees.

The results show that the proposed model performs better than the other models in terms of accuracy,

precision, and recall. The authors also analyse the contribution of different features to fraud detection and

discuss the limitations of the proposed model, such as the need for a large amount of data and the

challenges associated with handling imbalanced datasets.

1.3 “A Research Paper on Credit Card Fraud Detection”

The proposed model involves pre-processing the credit card transaction data and then apply- ing various

machine learning algorithms, such as Decision Trees, Random Forest, K-Nearest Neighbour, Naive Bayes,

and Artificial Neural Networks, to classify transactions as either fraudulent or non-fraudulent.

The authors evaluate the performance of their model using a dataset of credit card transactions and

compare it with other existing models, such as Logistic Regression, Support Vector Machine, and Gradient

Boosting Machine. The results show that the proposed model outperforms the other models in terms of

accuracy, precision, recall, and F1-score.

The paper also discusses the challenges associated with credit card fraud detection, such as the need for

real-time detection, the challenges of handling imbalanced datasets, and the importance of feature selection

for improving the performance of the model.

1.4 “Credit Card Fraud Detection Predictive Modelling”

The paper covers different machine learning techniques such as supervised, unsupervised, semi-supervised,

and deep learning, and how they are applied in credit card fraud detection. The authors provide a detailed

explanation of each technique, including its advantages and limitations, and also present a comparative

analysis of various machine learning techniques in terms of their performance metrics.

The paper also discusses the challenges associated with credit card fraud detection, such as imbalanced

datasets, the need for real-time detection, and the importance of feature selection for improving the

performance of the model.

The authors conclude the paper by highlighting the potential of machine learning techniques in credit card

fraud detection and the need for further research in this area.

1.5 “A machine learning based credit card fraud detection using the GA algorithm

for feature selection”

The paper covers different deep learning techniques, such as deep neural networks, convolutional neural

networks, recurrent neural networks, and auto encoders, and how they are applied in credit card fraud

detection. The authors provide a detailed explanation of each technique, including their advantages and

limitations.

The paper also presents a comprehensive analysis of different studies that have used deep learning

approaches for credit card fraud detection. The authors provide a detailed summary of each study,

including the dataset used, the deep learning technique applied, and the performance metrics obtained

1.1 “Credit Card Fraud Detection System”

The paper discusses the problem of credit card fraud and highlights the importance of detecting fraudulent

transactions promptly. The authors then present their proposed model, which involves pre-processing the

transaction data and then applying various machine learning algorithms, such as logistic regression,

decision trees, and random forests, to classify transactions as either fraudulent or non-fraudulent.

The authors evaluate the performance of their model using a dataset of credit card transactions and

compare it to other models, such as neural networks and support vector machines. The results show that

the proposed model performs better than the other models in terms of accuracy, precision, and recall.

The paper also discusses the limitations of the proposed model, such as the need for a large amount of data

and the challenges associated with handling imbalanced datasets. The authors conclude by highlighting the

potential of machine learning techniques in credit card fraud detection and the need for further research in

this area.

1.2 “Credit Card Fraud Detection using Machine Learning Algorithms”

The authors present their proposed model, which involves pre-processing the transaction data and then

applying a CNN for feature extraction and classification.

To improve the performance of the model, the authors also use feature selection techniques to identify the

most relevant features for fraud detection. They evaluate the performance of their model using a dataset of

credit card transactions and compare it to other models, such as logistic regression and decision trees.

The results show that the proposed model performs better than the other models in terms of accuracy,

precision, and recall. The authors also analyse the contribution of different features to fraud detection and

discuss the limitations of the proposed model, such as the need for a large amount of data and the

challenges associated with handling imbalanced datasets.

1.3 “A Research Paper on Credit Card Fraud Detection”

The proposed model involves pre-processing the credit card transaction data and then apply- ing various

machine learning algorithms, such as Decision Trees, Random Forest, K-Nearest Neighbour, Naive Bayes,

and Artificial Neural Networks, to classify transactions as either fraudulent or non-fraudulent.

The authors evaluate the performance of their model using a dataset of credit card transactions and

compare it with other existing models, such as Logistic Regression, Support Vector Machine, and Gradient

Boosting Machine. The results show that the proposed model outperforms the other models in terms of

accuracy, precision, recall, and F1-score.

The paper also discusses the challenges associated with credit card fraud detection, such as the need for

real-time detection, the challenges of handling imbalanced datasets, and the importance of feature selection

for improving the performance of the model.

1.4 “Credit Card Fraud Detection Predictive Modelling”

The paper covers different machine learning techniques such as supervised, unsupervised, semi-supervised,

and deep learning, and how they are applied in credit card fraud detection. The authors provide a detailed

explanation of each technique, including its advantages and limitations, and also present a comparative

analysis of various machine learning techniques in terms of their performance metrics.

The paper also discusses the challenges associated with credit card fraud detection, such as imbalanced

datasets, the need for real-time detection, and the importance of feature selection for improving the

performance of the model.

The authors conclude the paper by highlighting the potential of machine learning techniques in credit card

fraud detection and the need for further research in this area.

1.5 “A machine learning based credit card fraud detection using the GA algorithm

for feature selection”

The paper covers different deep learning techniques, such as deep neural networks, convolutional neural

networks, recurrent neural networks, and auto encoders, and how they are applied in credit card fraud

detection. The authors provide a detailed explanation of each technique, including their advantages and

limitations.

The paper also presents a comprehensive analysis of different studies that have used deep learning

approaches for credit card fraud detection. The authors provide a detailed summary of each study,

including the dataset used, the deep learning technique applied, and the performance metrics obtained

1.1 “Credit Card Fraud Detection System”

The paper discusses the problem of credit card fraud and highlights the importance of detecting fraudulent

transactions promptly. The authors then present their proposed model, which involves pre-processing the

transaction data and then applying various machine learning algorithms, such as logistic regression,

decision trees, and random forests, to classify transactions as either fraudulent or non-fraudulent.

The authors evaluate the performance of their model using a dataset of credit card transactions and

compare it to other models, such as neural networks and support vector machines. The results show that

the proposed model performs better than the other models in terms of accuracy, precision, and recall.

The paper also discusses the limitations of the proposed model, such as the need for a large amount of data

and the challenges associated with handling imbalanced datasets. The authors conclude by highlighting the

potential of machine learning techniques in credit card fraud detection and the need for further research in

this area.

1.2 “Credit Card Fraud Detection using Machine Learning Algorithms”

The authors present their proposed model, which involves pre-processing the transaction data and then

applying a CNN for feature extraction and classification.

To improve the performance of the model, the authors also use feature selection techniques to identify the

most relevant features for fraud detection. They evaluate the performance of their model using a dataset of

credit card transactions and compare it to other models, such as logistic regression and decision trees.

The results show that the proposed model performs better than the other models in terms of accuracy,

precision, and recall. The authors also analyse the contribution of different features to fraud detection and

discuss the limitations of the proposed model, such as the need for a large amount of data and the

challenges associated with handling imbalanced datasets.

1.3 “A Research Paper on Credit Card Fraud Detection”

The proposed model involves pre-processing the credit card transaction data and then apply- ing various

machine learning algorithms, such as Decision Trees, Random Forest, K-Nearest Neighbour, Naive Bayes,

and Artificial Neural Networks, to classify transactions as either fraudulent or non-fraudulent.

The authors evaluate the performance of their model using a dataset of credit card transactions and

compare it with other existing models, such as Logistic Regression, Support Vector Machine, and Gradient

Boosting Machine. The results show that the proposed model outperforms the other models in terms of

accuracy, precision, recall, and F1-score.

The paper also discusses the challenges associated with credit card fraud detection, such as the need for

real-time detection, the challenges of handling imbalanced datasets, and the importance of feature selection

for improving the performance of the model.

1.4 “Credit Card Fraud Detection Predictive Modelling”

The paper covers different machine learning techniques such as supervised, unsupervised, semi-supervised,

and deep learning, and how they are applied in credit card fraud detection. The authors provide a detailed

explanation of each technique, including its advantages and limitations, and also present a comparative

analysis of various machine learning techniques in terms of their performance metrics.

The paper also discusses the challenges associated with credit card fraud detection, such as imbalanced

datasets, the need for real-time detection, and the importance of feature selection for improving the

performance of the model.

The authors conclude the paper by highlighting the potential of machine learning techniques in credit card

fraud detection and the need for further research in this area.

1.5 “A machine learning based credit card fraud detection using the GA algorithm

for feature selection”

The paper covers different deep learning techniques, such as deep neural networks, convolutional neural

networks, recurrent neural networks, and auto encoders, and how they are applied in credit card fraud

detection. The authors provide a detailed explanation of each technique, including their advantages and

limitations.

The paper also presents a comprehensive analysis of different studies that have used deep learning

approaches for credit card fraud detection. The authors provide a detailed summary of each study,

including the dataset used, the deep learning technique applied, and the performance metrics obtained

* 1. “Credit Card Fraud Detection System”

The paper discusses the problem of credit card fraud and highlights the importance of detecting fraudulenttransactions promptly. The authors then present their proposed model, which involves pre-processing thetransaction data and then applying various machine learning algorithms, such as logistic regression,decision trees, and random forests, to classify transactions as either fraudulent or non-fraudulent.The authors evaluate the performance of their model using a dataset of credit card transactions andcompare it to other models, such as neural networks and support vector machines. The results show thatthe proposed model performs better than the other models in terms of accuracy, precision, and recall.The paper also discusses the limitations of the proposed model, such as the need for a large amount of dataand the challenges associated with handling imbalanced datasets. The authors conclude by highlighting thepotential of machine learning techniques in credit card fraud detection and the need for further research inthis area.

1.2 “Credit Card Fraud Detection using Machine Learning Algorithms”

The authors present their proposed model, which involves pre-processing the transaction data and thenapplying a CNN for feature extraction and classification.To improve the performance of the model, the authors also use feature selection techniques to identify themost relevant features for fraud detection. They evaluate the performance of their model using a dataset ofcredit card transactions and compare it to other models, such as logistic regression and decision trees.The results show that the proposed model performs better than the other models in terms of accuracy,precision, and recall. The authors also analyse the contribution of different features to fraud detection anddiscuss the limitations of the proposed model, such as the need for a large amount of data and thechallenges associated with handling imbalanced datasets.

1.3 “A Research Paper on Credit Card Fraud Detection”

The proposed model involves pre-processing the credit card transaction data and then apply- ing variousmachine learning algorithms, such as Decision Trees, Random Forest, K-Nearest Neighbour, Naive Bayes,and Artificial Neural Networks, to classify transactions as either fraudulent or non-fraudulent.The authors evaluate the performance of their model using a dataset of credit card transactions andcompare it with other existing models, such as Logistic Regression, Support Vector Machine, and GradientBoosting Machine. The results show that the proposed model outperforms the other models in terms ofaccuracy, precision, recall, and F1-score.The paper also discusses the challenges associated with credit card fraud detection, such as the need forreal-time detection, the challenges of handling imbalanced datasets, and the importance of feature selectionfor improving the performance of the model.

1.4 “Credit Card Fraud Detection Predictive Modelling”

The paper covers different machine learning techniques such as supervised, unsupervised, semi-supervised,and deep learning, and how they are applied in credit card fraud detection. The authors provide a detailedexplanation of each technique, including its advantages and limitations, and also present a comparativeanalysis of various machine learning techniques in terms of their performance metrics.The paper also discusses the challenges associated with credit card fraud detection, such as imbalanceddatasets, the need for real-time detection, and the importance of feature selection for improving theperformance of the model.The authors conclude the paper by highlighting the potential of machine learning techniques in credit cardfraud detection and the need for further research in this area.

1.5 “A machine learning based credit card fraud detection using the GA algorithmfor feature selection”

The paper covers different deep learning techniques, such as deep neural networks, convolutional neuralnetworks, recurrent neural networks, and auto encoders, and how they are applied in credit card frauddetection. The authors provide a detailed explanation of each technique, including their advantages andlimitations.The paper also presents a comprehensive analysis of different studies that have used deep learningapproaches for credit card fraud detection. The authors provide a detailed summary of each study,including the dataset used, the deep learning technique applied, and the performance metrics obtained

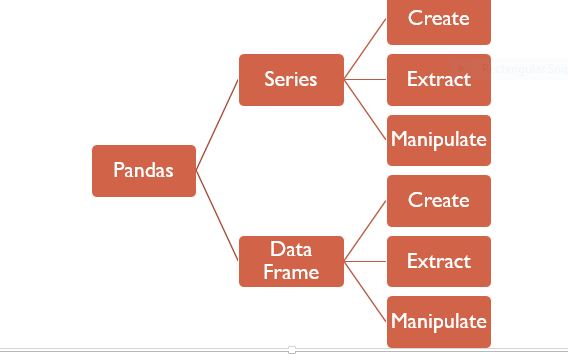
Methodology

Overview of Pandas

* Pandas is derived from the term "panel data", an econometrics term for data sets that include observations over multiple time periods for the same individuals
* Data in pandas is used to feed statistical analysis in **SciPy**, plotting functions from **Matplotlib**, and machine learning algorithms in **Scikit-learn**.
* The primary two components of pandas are the Series and Data Frame.

What can you do with pandas?

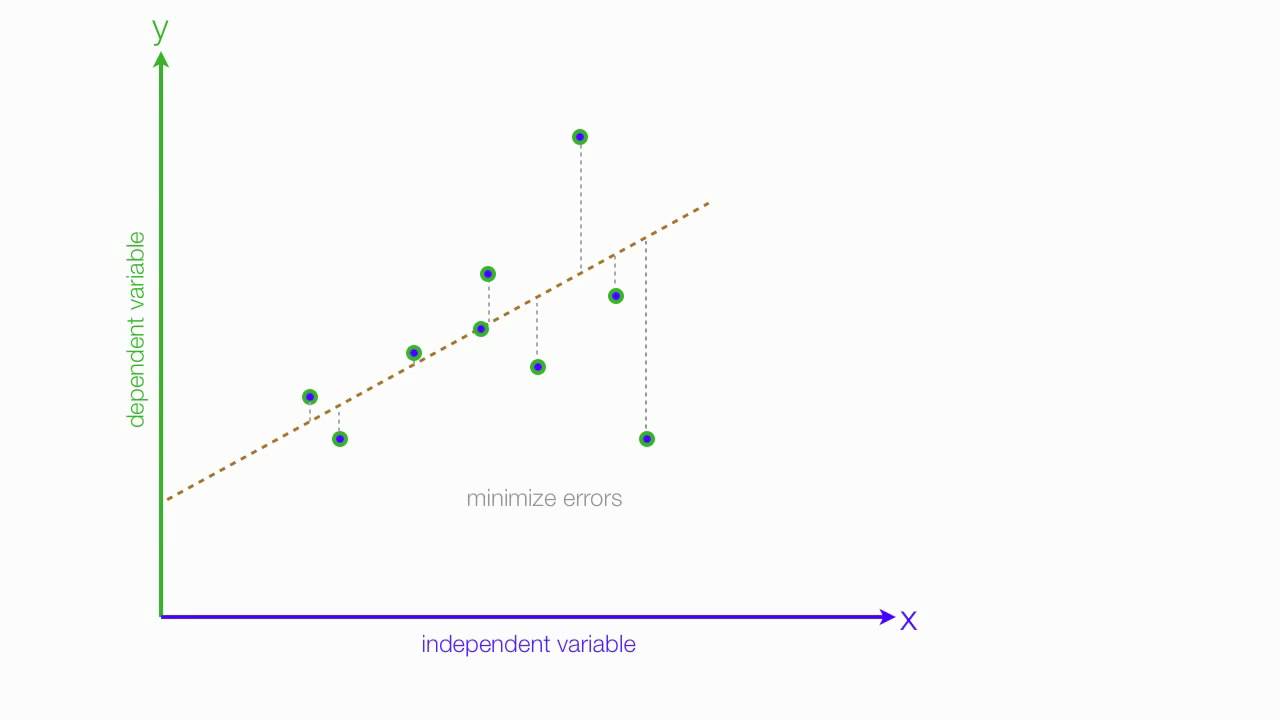
* Easily calculate statistics about data such as finding the average, distribution, and median of columns
* Use data visualization tools, such as Matplotlib, to easily create plot bars, histograms, and more
* Clean your data by filtering columns by particular criteria or easily removing values
* Manipulate your data flexibly using operations like merging, joining,

reshaping, and more

* Read, write, and store your clean data as a txt file, json file or csv file etc.

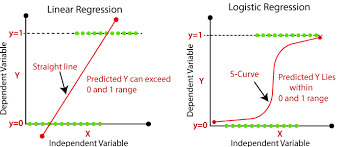
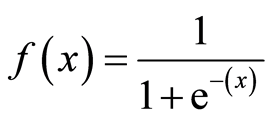
Linear Regression

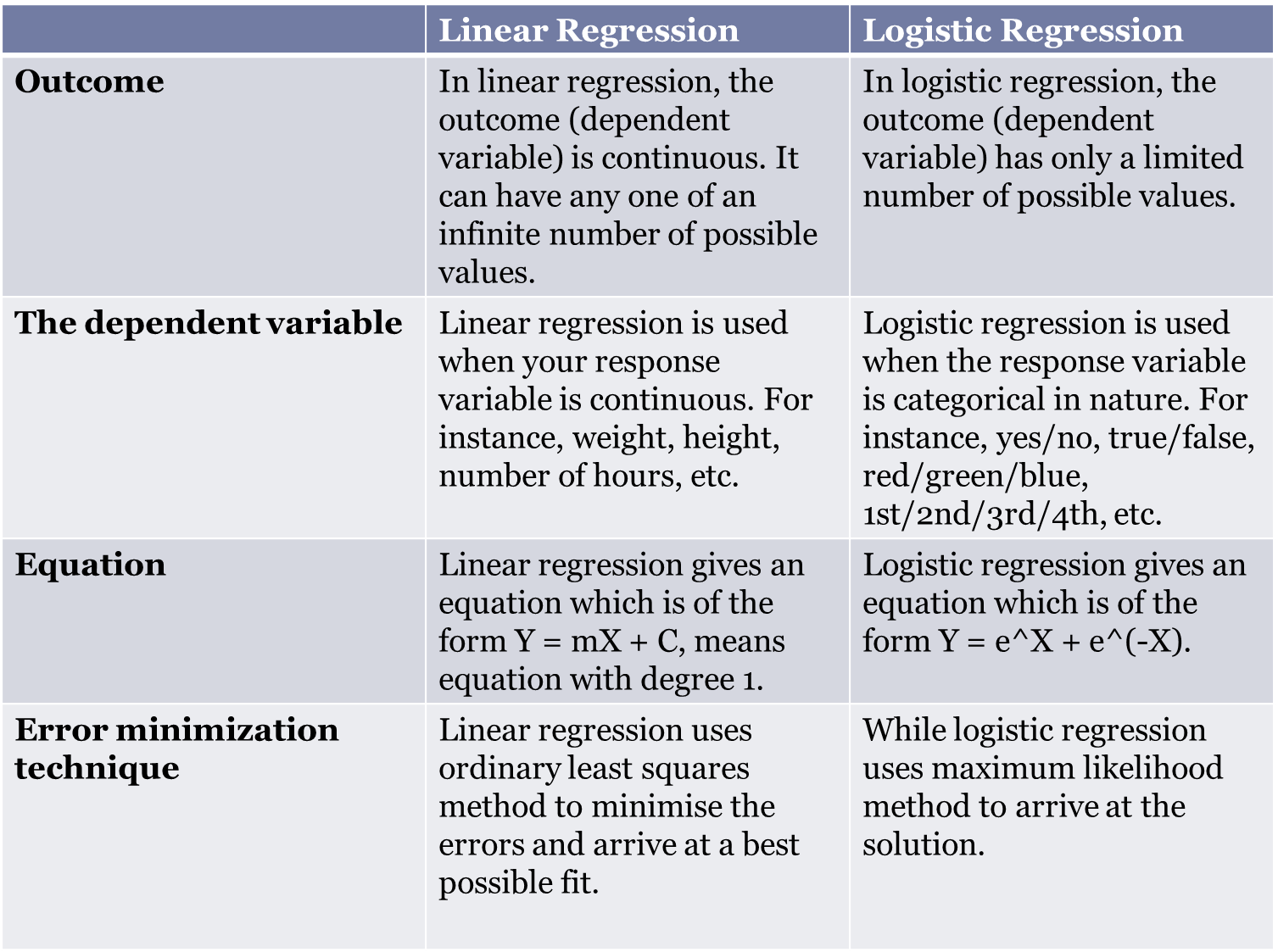
* Linear Regression is the simplest supervised algorithm for regression problems.
* Regression problems usually have ‘continuous’ dependent features.
* The dependent features are called the dependent variables, outputs, or responses.
* The independent features are called the independent variables, inputs, or predictors.
* Applications in Price predictions, Score predictions, GDP etc.
* Graphical Representation



Logistic regression

* Logistic regression is a supervised , classification algorithm used to assign observations to a discrete set of classes.
* Examples of classification problems are : Email spam or not spam, Online transactions Fraud or not Fraud, Tumour Malignant or Benign
* Logistic regression transforms its output using the logistic sigmoid function to return a probability value
* It’s an S-shaped curve that can take any real-valued number and map it into a value between 0 and 1





Conclusion

Credit card fraud is a persistent problem that can lead to significant financial losses for individuals andbusinesses alike. With the increasing reliance on electronic payments, detecting and preventing fraud hasbecome a crucial task for financial institutions.In recent years, various techniques and algorithms have been developed to improve the accuracy andefficiency of credit card fraud detection. These techniques include rule-based systems, statistical methods,machine learning, and deep learning.Machine learning algorithms, in particular, have shown promising results in detecting credit card fraud, asthey can learn from large datasets and identify patterns that are difficult for human analysts to detect. Deeplearning techniques, such as neural networks, have also shown great potential for detecting fraudulenttransactions.However, credit card fraudsters continue to develop new and sophisticated methods to evadedetection, and fraud detection systems must continue to evolve and adapt to these changing threats.

References

1. “Credit Card Fraud Detection System” by K. Madkaikar, M. Nagvekar, P. Parab, R. Raikar, S. Patil

2. .“Credit Card Fraud Detection using Machine Learning Algorithms” by V. Dornadulaa, Geetha S.

3. “A Research Paper on Credit Card Fraud Detection” by S. Teja, B. Munendra, S. Gokulkrishnan

4. “Credit Card Fraud Detection Predictive Modelling” by N. Sharma

5. “A machine learning based credit card fraud detection using the GA algorithm for feature selection” by

E. Ileberi, Y. Sun, Z. Wang

6. [1] Dataset: http://packages.revolutionanalytics.com/datasets/

7. Jain R., Gour B., Dubey S., A hybrid approach for credit card fraud detection using rough set and

decision tree technique, International Journal of Computer Applications 139(10) (2016).

8. Credit, ,Card, , Fraud, , Detection, , Based,,on,,Transaction, , Behavior, , -by,John,Richard,D., , , ,

Kho,, , Larry,A.,Vea”,published,by,Proc.,of, , the,2017, , IEEE, , Region,,10,,Conference,

(TENCON),, , Malaysia,, , November,,5-8,,2017

9. An Intelligent Approach to Credit Card Fraud Detection Using an Optimised Light Gradient Boosting

Machine ALTYEB ALTAHER TAHA AND SHARAF JAMEEL MALEBAR.

10. https://www.kaggle.com/mlg-ulb/creditcardfraud

11. Campus K. Credit card fraud detection using machine learning models and collating machine learning

models. Int J

Pure Appl Math. 2018;118(20):825–38.

12. [5]Credit card fraud [Online]

1. “Credit Card Fraud Detection System” by K. Madkaikar, M. Nagvekar, P. Parab, R. Raikar, S. Patil

2. .“Credit Card Fraud Detection using Machine Learning Algorithms” by V. Dornadulaa, Geetha S.

3. “A Research Paper on Credit Card Fraud Detection” by S. Teja, B. Munendra, S. Gokulkrishnan

4. “Credit Card Fraud Detection Predictive Modelling” by N. Sharma

5. “A machine learning based credit card fraud detection using the GA algorithm for feature selection” byE. Ileberi, Y. Sun, Z. Wang

6. [1] Dataset: <http://packages.revolutionanalytics.com/datasets/>

7. Jain R., Gour B., Dubey S., A hybrid approach for credit card fraud detection using rough set anddecision tree technique, International Journal of Computer Applications 139(10) (2016).

8. Credit, ,Card, , Fraud, , Detection, , Based,,on,,Transaction, , Behavior, , -by,John,Richard,D., , , ,Kho,, , Larry,A.,Vea”,published,by,Proc.,of, , the,2017, , IEEE, , Region,,10,,Conference,(TENCON),, , Malaysia,, , November,,5-8,,2017

9. An Intelligent Approach to Credit Card Fraud Detection Using an Optimised Light Gradient BoostingMachine ALTYEB ALTAHER TAHA AND SHARAF JAMEEL MALEBAR.

10. <https://www.kaggle.com/mlg-ulb/creditcardfraud>

11. Campus K. Credit card fraud detection using machine learning models and collating machine learningmodels. Int JPure Appl Math. 2018;118(20):825–38.

12. [5]Credit card fraud [Online] <https://www.kaggle.com/mlg-ulb/creditcardfraud>