*A*

***PROJECT***

***Synopsis***

***On***

# “Detecting Credit Card Fraud Using Machine Learning”

*Submitted*

*in partial fulfilment*

*for the award of the degree of*

***Bachelor of Technology***

in Information Technology



**Submitted to:**  **Submitted by:**

Dr. Sunil Kumar Jangir Asmita Goswami(15EJCIT014)

Head of Department Lokesh Soni (15EJCIT043)

Harsh Banshiwal(15EJCIT030)

Anshul Jain(15EJCIT009)

**DEPARTMENT OF INFORMATION TECHNOLOGY**

**JAIPUR ENGINEERING COLLEGE & RESEARCH CENTRE**

Rajasthan Technical University, Kota

**(Session 2018-2019)**

**CONTENT**

|  |  |  |
| --- | --- | --- |
| **S.NO.** | **Table Name** | **Page No.** |
| 1 | Introduction | 3 |
| 2 | Objective | 3 |
| 3 | Methodology | 4 |
| 4 | Tool Used | 5 |
| 5 | Advantage and Disadvantage of the Project | 6 |
| 6 | Application of the Project | 7 |
| 7 | References | 7 |

**MAJOR Project Synopsis**

**TITLE:** Detecting Credit Card Fraud Using Machine Learning

**Location:**Department of Information Technology

(JECRC, Jaipur), Rajasthan Technical University.

**1. Introduction**

For years, fraudsters would simply take numbers from **credit or debit card** sand print them onto blank plastic cards to use at brick-and-mortar stores. But in 2015, **Visa**and**Mastercard**mandated that banks and merchants introduce **EMV —**chip card technology, which made it possible for merchants to start requesting a PIN for each transaction. Nevertheless, experts predict online credit card fraud to soar to a whopping **$32 billion**in 2020.

Putting it into perspective, this amount is superior to the profits posted recently by some worldwide household, blue chip companies in 2017, such as **Coca-Cola**($2 billions), Warren Buffet’s **Berkshire Hathaway** ($24 billions) and **JP Morgan Chase** ($23.5 billions).In addition to the implementation of chip card technology, companies have been investing massive amounts in other technologies for detecting fraudulent transactions.

**2. Objective**

The Objective behind this is to demonstrate the technical feasibility of a deep learning approach to design of an effective fraud detection system is necessary in order to reduce the losses incurred by the customers and financial companies. Research has been done on many models and methods to prevent and detect frauds.

**3. Methodology/ Planning of work:**

In **Machine Learning**, problems like fraud detection are usually framed as **classification problems —**predicting a discrete class label output given a data observation. Examples of classification problems that can be thought of are **Spam Detectors, Recommender Systems**and**Loan Default Prediction.**

Talking about the credit card payment fraud detection, the classification problem involves creating models that have enough intelligence in order to properly classify transactions as either **legit** or **fraudulent**, based on transaction details such as **amount, merchant, location, time**and others.

Financial fraud still amounts for considerable amounts of money. Hackers and crooks around the world are always looking into new ways of committing financial fraud at each minute. Relying exclusively on rule-based, conventionally programmed systems for detecting financial fraud would not provide the appropriate time-to-market. This is where **Machine Learning** shines as a unique solution for this type of problem.

The main challenge when it comes to modelling fraud detection as a classification problem comes from the fact that in real world data, the majority of transactions is not fraudulent. Investment in technology for fraud detection has increased over the years so this shouldn’t be a surprise, but this brings us a problem: **imbalanced data**.

**4. Tool Used:**

**4.1 Jupyter Notebook:**

The Jupyter Notebook is an open-source web application that allows us to create and share documents that contain live code, equations, visualizations and narrative text. Uses include data cleaning and transformation, numerical simulation, statistical modelling, data visualization, machine learning, and much more.

**4.2 Anaconda Distribution:**

Conda is a powerful package manager and environment manager that you use with command line commands at the Anaconda Prompt for Windows, or in a Terminal window for macOS or Linux.

**4.3 Code Libraries Used:** Keras (keras.ai), Numpy, Pandas.

**5. Advantage and Disadvantage of The Project:**

* Fraud detection is a complex issue that requires a substantial amount of planning before throwing machine learning algorithms at it. Nonetheless, it is also an application of data science and machine learning for the good, which makes sure that the customer’s money is safe and not easily tampered with.

**6. Application of The Project:**

Given the data project will detect the credit card fraud.

**7. References**

[1]Nick Becker — The Right Way to Oversample in Predictive Modelling

[2]Andrea Dal Pozzolo, Olivier Caelen, Reid A. Johnson and Gianluca Bontempi. *Calibrating Probability with Undersampling for Unbalanced Classification*. In Symposium on Computational Intelligence and Data Mining (CIDM), IEEE, 2015

[3] L.J.P. van der Maaten and G.E. Hinton, Visualizing High-Dimensional Data Using t-SNE (2014), Journal of Machine Learning Research

[4] Machine Learning Group — ULB, Credit Card Fraud Detection (2018), Kaggle

[5] Nathalie Japkowicz, Learning from Imbalanced Data Sets: A Comparison of Various Strategies (2000), AAAI Technical Report WS-00–05