|  |  |
| --- | --- |
|  | **Sri Lanka Institute of Information Technology** |

PROJECT REGISTRATION FORM

|  |
| --- |
|  |

(This form should be completed and submitted on or before 3.00 PM, Friday 3rd March, 2017)

The purpose of this form is to allow final year students of the B.Sc. (Hon) degree program to enlist in the final year project group. Enlisting in a project entails specifying the project title and the details of four members in the group, the internal supervisor (compulsory), external supervisor (may be from the industry) and indicating a brief description of the project. The description of the project entered on this form will not be considered as the formal project proposal. It should however indicate the scope of the project and provide the main potential outcome.

|  |  |
| --- | --- |
| PROJECT TITLE | Credit card fraud detection |

|  |  |
| --- | --- |
| RESEARCH GROUP | Data Science, Machine Learning |

|  |  |  |
| --- | --- | --- |
| PROJECT NUMBER |  | (will be assigned by the lecture in charge) |

PROJECT GROUP MEMBER DETAILS: (Please start with group leader’s details)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | STUDENT NAME | STUDENT NO. | CONTACT NO. | EMAIL ADDRESS |
| 1 | T.M.G.A.B. Thennakoon  (GROUP LEADER) | IT15046512 | 0713761687 | [anuruddha.thennakoon@gmail.com](mailto:anuruddha.thennakoon@gmail.com) |
| 2 | H.G.S. Premadasa | IT15142610 | 0718717337 | [sasithasa@gmail.com](mailto:sasithasa@gmail.com) |
| 3 | C.B.P. Lochana | IT15111784 | 0714588990 | [cheebhagyani1994@gmail.com](mailto:cheebhagyani1994@gmail.com) |
| 4 | M.D.S. Mihiranga | IT15004550 | 0702609083 | [shalitha.mihi99@gmail.com](mailto:shalitha.mihi99@gmail.com) |

|  |  |  |
| --- | --- | --- |
| SUPERVISOR | | |
| Mr. Nuwan Kuruwitaarachchi |  |  |
| Name | Signature | Date |

|  |  |  |
| --- | --- | --- |
| CO-SUPERVISOR (will be assigned by the Supervisor, if necessary) | | |
|  |  |  |
| Name | Signature | Date |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| EXTERNAL SUPERVISOR (if any, may be from the industry) | | | | |
|  |  |  |  |  |
| Name | Affiliation | Contact Address | Contact Numbers | Signature/Date |

|  |  |  |
| --- | --- | --- |
| ACCEPTANCE BY CDAP MEMBER | | |
|  |  |  |
| Name | Signature | Date |

PROJECT DETAILS

|  |
| --- |
| Brief Description of your Research Problem: |
| This research project is based on fraud detection, associated with the most popular mode of payment, which is credit card. Credit card payments are used for both online and regular purchases and with the increase of their usage the probability of happening a fraud has risen.  Credit card frauds are financial losses to banks. It’s a must to analyze those transactions and detect the frauds which causes big losses.  We can categorize credit card fraud as follows.   1. Theft Fraud/Counterfeit Fraud - In case of Theft Fraud the Credit Card is stolen by an unknown person or the lost card is found by someone and it is used by them many times as possible until the card is blocked by the bank. In case of the Counterfeit fraud the Card is used remotely by using the card detail of the card holder. 2. Application Fraud - In Application fraud the fraudster applies with false documents or with stolen documents of someone else to get a Credit Card. 3. Behavioral Fraud - Behavioral fraud occurs when details of legitimate cards have been obtained fraudulently and sales are made on a ’card holder’ present basis. 4. Cramming/Salami Fraud - Cramming is when a fraudster using a Credit Card makes piecemeal transactions over a long period of time. 5. Triangulation Fraud - The fraudster maintains some web site through which the items are offered pay on delivery at very high discount.   Here we can list down some models that have been developed by researchers so far and the techniques that have been used.   * **Dempster–Shafer theory and Bayesian learning** - Four components of Dempster Shafer Theory are: rule based filter, Dempster–Shafer adder, transaction history database and Bayesian learner. * **Hidden Markov Model** - Fraud detection system receives each incoming transaction and check the purchasing behavior for finding whether the transaction is fraud or not. * **Detection of Credit Card Fraud using Fuzzy** - Fuzzy Darwinian Detection system use fuzzy logic rules. The system consists of a fuzzy expert system and a Genetic Programming (GP) search algorithm. * **Bayesian and Neural Networks** - Bayesian networks or belief networks use machine learning and data mining methods of artificial intelligence. * **Fraud detection using Association rules** - In this fuzzy logic is used with association rules to overcome the difficulties of minimum support and confidence, optimizes the execution time. * **SVM and Logistic regression** - This paper examines the performance of two data mining techniques Support vector machine and random forest with the logistic regression in credit card fraud detection. * **Fraud detection using Decision Tree and SVM** - Decision tree model built in this study use three classification models C5.0, C&RT and SVM and CHAID. * **Max entropy with Bayesian learning** - Credit card fraud detection system is proposed which detect frauds using maximum entropy and Bayesian learning. * **Fraud detection using Rule based expert system** - A rule based expert is proposed for detecting fraud in credit card. * **Fraud detection using Peer Group analysis** - A peer group is a group of records which behave same as the target records. When the behavior of the target record is different from its peers, an anomaly can be signed.   Techniques used by the above mentioned models.   * Descriptive (Unsupervised)   + - * Graphical Outlier Detection Procedures       * Statistical Outlier Detection Procedures       * Break-Point Analysis       * Peer-Group Analysis       * Association Rule Analysis       * Clustering * Predictive (Supervised) * Regression. * Linear Regression * Logistic Regression * Classification   + Decision Tree   + Neural Network   + Support Vector Machine(SVM)   + Naive Baye’s Rule based classification * Every model has shown that it works with a unique dataset, but not any model is appropriate for all datasets. * With the development of the technology, the accuracy and the efficiency of the detected fraud is important. |
| Description of the Solution:   * Selecting the most suitable technique from the above used techniques, detect fraudulent behavior in a credit card transaction by providing a trained learner with data by using both supervised and unsupervised techniques which improves accuracy and efficiency. * Then the result is visualized in a 2D format (histogram/ bar chart) |

|  |
| --- |
| Main expected outcomes of the project: |
| * The implemented program should be able to classify new transactions. * Based on previous data can recognize unusual patterns in any payment procedure. * Drastic reduction of number of false positives transactions identified as fraudulent transactions although they are normal transactions and increase of efficiency in detection process. |

WORKLOAD ALLOCATION (Please provide a brief description about the workload allocation)

|  |  |
| --- | --- |
| MEMBER 1 | ………………………………………………………………………………………………………………………………………………………… |
| * Data preprocessing.   Ex :-   * Resampling the dataset to have an approximate 50-50 ratio. One way to achieve this is by oversampling another is undersampling. | |

|  |  |
| --- | --- |
| MEMBER 2 | ………………………………………………………………………………………………………………………………………… |
| * Use a Predictive(Supervised) learning technique to detect credit card fraud for the above preprocessed data.   Ex :-   * Can use widely used Neural networks technique. | |

|  |  |
| --- | --- |
| MEMBER 3 | ………………………………………………………………………………………………………………………………………… |
| * Use a Descriptive(Unsupervised) learning technique to detect credit card fraud for the above preprocessed data   Ex :-   * Can use widely used Clustering technique. | |

|  |  |
| --- | --- |
| MEMBER 4 | ………………………………………………………………………………………………………………………………………… |
| * Visualization of the result in a graphical view for the better understanding.   Ex :-   * Visualization on a 2D map with t-SNE | |

DECLARATION

“We declare that the project would involve material prepared by the Group members and that it would not fully or partially incorporate any material prepared by other persons for a fee or free of charge or that it would include material previously submitted by a candidate for a Degree or Diploma in any other University or Institute of Higher Learning and that, to the best of our knowledge and belief, it would not incorporate any material previously published or written by another person in relation to another project except with prior written approval from the supervisor and/or the coordinator of such project and that such unauthorized reproductions will construe offences punishable under the SLIIT Regulations.

We are aware, that if we are found guilty for the above mentioned offences or any project related plagiarism, the SLIIT has right to suspend the project at any time and or to suspend us from the examination and or from the Institution for minimum period of one year”.

|  |  |  |  |
| --- | --- | --- | --- |
|  | STUDENT NAME | STUDENT NO. | SIGNATURE |
| 1 | T.M.G.A.B. Thennakoon  (GROUP LEADER) | IT15046512 |  |
| 2 | H.G.S. Premadasa | IT15142610 |  |
| 3 | C.B.P. Lochana | IT15111784 |  |
| 4 | M.D.S. Mihiranga | IT15004550 |  |