

**Overview of Industry Project for Students**

The objectives of an Industry Project are to:

* Expose students to the challenges, realities and considerations while working on an IT project within the industry.
* Provide opportunities for students to interact with industry practitioners and obtain feedback on their produced work.
* Allow students to deliver outcomes and value that companies can potentially apply and implement within their business or operating environment.

**Contact Information**

|  |  |
| --- | --- |
| Company / Organisation Details: | D-Risk Technology  (UEN: 201725582D) |
| Contact Person Details (including name, email, designation, contact number, etc.): | Mr. SEE Weng Lung  CEO  wlsee@d-risk.tech  (Do **NOT** contact unless your supervisor gives permission) |

**Project Proposal**

|  |  |
| --- | --- |
| Proposed Project Title  (The title may be amended by SOI for internal use) | Credit Risk Analysis using Machine Learning |
| Project Objective and Motivation  (Please state the purpose of the project as well as the importance to your company  or organisation) | To understand the workings of an industry required application that performs credit risk analysis for small and medium enterprises through machine learning, and provide meaningful explanation via explainable AI (XAI). |
| Project Description and Details  (This section should include details regarding the project scope, project user and/or technical requirements, and/or any other information that would be useful in understanding the project) | **Background**  In lending, credit risk refers to the potential risk a lender faces when providing loans to a borrower. To mitigate this risk, lenders analyze various factors related to a borrower’s creditworthiness, such as their current debt load and income.  Several companies excel in performing credit risk assessments, including Standard & Poor’s (S&P), Moody’s, and Fitch. Banks, in particular, conduct extensive credit assessments before approving loans to corporations. Highly skilled credit analysts, with deep understanding of industries and finance, spend significant time conducting both quantitative and qualitative analyses on borrowing corporations to determine their creditworthiness. One of the best indicators of creditworthiness is a credit rating. However, credit ratings performed by global credit rating agencies such as Standard & Poor’s are expensive and can take more than four weeks to complete.  As a result, lenders like banks have internal credit departments to perform credit assessments. However, this manual process is time-consuming, requires high skill levels, and can be subjective. Small and medium-sized enterprise (SME) loans, which are typically much smaller, require the same resources as larger corporate loans. Consequently, small loans do not provide sufficient profitability for banks to process, leading to difficulties for SMEs in obtaining financing.  D-Risk believes that the solution to SME lending challenges lies in the availability of technology that is easy to use, quick, and affordable for assessing the creditworthiness of SMEs. To this end, D-Risk has developed an AI-driven credit rating model to evaluate SMEs. However, while AI models can predict credit ratings, they often lack transparency in their decision-making processes. Therefore, D-Risk aims to explore the use of Explainable AI (XAI) to provide insights into how the model arrives at its conclusions.  **FYP Proposal**  **Objective:** D-Risk Technology aims to provide students with an opportunity to develop competencies in machine learning and explainable AI (XAI) by performing Credit Risk analysis using synthetic data. This project will be conducted over 13 weeks in a team of 4 students.  **Assumptions:** Students should have a basic understanding of Python and Machine Learning. If not, they are encouraged to use resources like Kaggle Learn to acquire the necessary skills.  **Dataset:** “Set B Corporate Rating.csv”  **Project Structure:**  **Low Code Phase:**   1. **Identifying the Best ML Model:**    * Use Knime to build workflows and identify the most accurate Machine Learning model for Credit Risk Analysis. 2. **Implementing XAI:**    * Extend the Knime workflow to include Explainable AI components.   **Python Coding Phase:**   1. **Identifying the Best ML Model:**    * Use Visual Studio Code or Jupyter Notebook to replicate the Knime workflows in Python. 2. **Implementing XAI:**    * Continue building Python code to implement XAI.   **Timeline and Deliverables:**  **Weeks 1-3:**   * **Meet with Mr. See Weng Lung:** Understand project requirements, dataset, and industry practices. * **Deliverables:**   + Document learning using the CRISP-DM method.   + Research and document credit risk and dataset headers (e.g., current ratio).   Some suggested Self study resources (may not be limited to these resources)  [Financial Statements: List of Types and How to Read Them (investopedia.com)](https://www.investopedia.com/terms/f/financial-statements.asp)  [Financial Ratios used in Credit Analysis | CFA Level 1 - AnalystPrep](https://analystprep.com/cfa-level-1-exam/fixed-income/financial-ratios-credit-analysis/)  [Corporate Credit Rating: What it is, How it Works (investopedia.com)](https://www.investopedia.com/terms/c/corporate-credit-rating.asp)  [Credit Rating: Definition and Importance to Investors (investopedia.com)](https://www.investopedia.com/terms/c/creditrating.asp)  Weeks 4-6:  Build the pipeline in Knime to train and test.  Student A: Logistic Regression  Student B: Random Forest  Student C: XGBoost  Student D: Decision Tree  Optional models (Supervisor’s Discretion):   * Neural Network * Quadratic Discriminat * Naïve Bayes * KNN * Gradient Boosting * Etc   Week 7 (Mid point check)  Gather feedback from Mr See Weng Lung and FYP Evaluators.  Term break – Week 9:  **Python Coding:**   * Use Visual Studio Code or Jupyter Notebook to replicate Knime workflows.   *Self-Study Resources*  [Learn Intro to Machine Learning | Kaggle](https://www.kaggle.com/learn/intro-to-machine-learning)  [Learn Intermediate Machine Learning | Kaggle](https://www.kaggle.com/learn/intermediate-machine-learning)  [Learn Machine Learning Explainability | Kaggle](https://www.kaggle.com/learn/machine-learning-explainability)  Week 10-11   * **Document Learning:** Achieve proficiency in explaining the models and their outcomes.   **Student Expectations:**   * Demonstrate proficiency in machine learning concepts and Python coding. * Take charge of individual models and perform cross-comparisons. * Explain model outcomes proficiently. * Compare outcomes of different models (e.g., XGBoost vs. Gradient Boosting Classifier).   **Possible Alternatives (Supervisor’s Discretion):**   * **Web Application:**   + **Admin Functions:**     - CRUD User     - CRUD Company     - CRUD company financials       * Include some calculation of financial ratios with numbers derived from financial statements.     - Upload financial statements (.pdf and .docx).     - Enter values for analysis fields.     - Export data.     - View predicted credit rating.     - Status management (draft, approved, expired).   + **Reports:**     - Admin View companies by risk category.     - Admin View top 10 best/worst companies.   + **Integration:**     - Integrate with Machine Learning and XAI. |
| Project Deliverable  (Please list or state the deliverables that is expected from the project) | A project as described above. |