Computer Science Capstone Topic Approval Form

The purpose of this document is to help you clearly explain your capstone topic, project scope, and timeline. Identify each of these areas so that you will have a complete and realistic overview of your project. Your instructor cannot sign off on your project topic without this information.

Note: You must fill out and submit this form. Space beneath each number will expand as needed.

Note: Any costs associated with developing the application will be the responsibility of the student.

INFORM INSTRUCTOR:
Potential use of proprietary company information: (Y/N)

ANALYSIS:

1. Project topic and description:

My project topic is a classification problem in loan application processing. I aim to create a machine learning algorithm trained on historical loan data to predict whether a loan applicant will default on a loan issued with the requested loan terms (loan amount, interest rate, term length, etc.).

The client for this project is the fictitious company "Origination Solutions Inc." This company provides a web-based loan origination platform for various lenders. Origination Solutions is interested in integrating a machine learning model into its platform to perform automated decision-making when processing loan applications. They have requested a prototype model capable of ingesting borrower information (debt-to-income ratio, credit score, etc.) and loan terms (loan amount, interest rate, etc.) to predict if the potential borrower will default on the requested loan. I will fill the role of the software engineer tasked with the design and implementation of this prototype.

Due to its efficiency and hands-off nature, this project will be an asset to Origination Solutions Inc. and its clients. The model will contribute to a reduced need for manual decision-making in loan origination, and the process of submitting loan data to the model for prediction will be much faster than the manual decision-making process.

2. Project purpose and goals:

This project aims to demonstrate my ability to design and build a machine learning-based loan application decision tool. Automated loan decision-making is an essential application of machine learning in the financial services industry, providing faster turnaround times and increased consistency compared to traditional manual review processes. This project will

WESTERN GOVERNORS UNIVERSITY

showcase my ability to build a machine learning model that can input borrower and loan information and output a reliable, accurate risk prediction to support automated loan approvals or rejections. The system will highlight my skills in data preparation, model training, and integration of predictive analytics into a business context.

3. Descriptive method:

My project will use data visualizations to describe and better understand the historical loan data before model training. The three visualizations I plan to create are:

A Distribution Plot: A histogram or density plot showing the distribution of borrower credit scores to understand the typical credit profile of applicants.

A Correlation Heatmap: A heatmap visualization of the correlations between key numeric variables such as debt-to-income ratio, loan amount, interest rate, and credit score to identify meaningful relationships.

A Feature Importance Plot: After initial model training, I will create a bar plot displaying the relative importance of each feature in predicting loan default risk, helping to interpret which borrower characteristics are most influential.

4. Predictive or prescriptive method:

I will use a Random Forest classifier as the predictive method for this project. Random Forest is an ensemble learning method that builds multiple decision trees and combines their outputs to improve predictive accuracy and control overfitting. After training the model on historical loan performance data, it will be able to predict the likelihood of default for new loan applicants based on the information they provide. By analyzing the output probability, Origination Solutions can make informed, automated decisions about loan approvals with greater speed and accuracy than manual methods.

DESIGN and DEVELOPMENT:

1. Computer science application type (select one):
   - Mobile (indicate Apple or Android)
   - <mark>Web</mark>
   - Stand-alone

2. Programming/development language(s) you will use:

I will use Python 3.11+ as the primary programming language. I also plan to use data analytics tools like Numpy and Matplotlib for data analysis. Scikit-learn will be used to implement the Random Forest classifier. The machine learning algorithm will run on the web application's backend; the Flask library will be used to provide ML services via API requests. Supporting development languages will include JavaScript and HTML for building the frontend user interface and analytics dashboard.

WESTERN GOVERNORS UNIVERSITY

3. Operating system(s) or platform(s) you will use:

Windows 10

4. Database Management System you will use:

N/A

5. Estimated number of hours for the following:
    i.   Planning and design: 20
    ii. Development: 40
    iii. Documentation: 20
    iv. Total:80

6. Projected completion date: May 31, 2025

IMPLEMENTATION and EVALUATION:

1. Describe how you will approach the execution of your project.
    a. Historical loan data will be retrieved from the LendingClub loan dataset available on Kaggle.
    b. The dataset will be analyzed, cleaned, and preprocessed.
    c. The model will be created and trained using the preprocessed data.
    d. Model accuracy and performance will be evaluated.
    e. If results are undesirable, steps B-D will be repeated.
    f. A Flask endpoint will be created for model prediction requests.
    g. A simple frontend will be created with HTML and JavaScript to provide a UI for model prediction requests.
    h. The project will be documented, and visualizations will be created.

✓ **This project does not involve human subjects research and is exempt from WGU IRB review.**

STUDENT'S SIGNATURE

_____**Jaxon Adams**_____

By signing and submitting this form, you acknowledge that any costs associated with the development and execution of the application will be your (the student's) responsibility.

WESTERN GOVERNORS UNIVERSITY

Charles Paddock

INSTRUCTOR'S SIGNATURE:

_____

INSTRUCTOR APPROVAL DATE:4/28/2025

WESTERN GOVERNORS UNIVERSITY.