**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 course 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.*

*Any cost associated with developing the application will be the responsibility of the student.*

**INFORM INSTRUCTOR:**

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

N

**ANALYSIS:**

1. **Project topic AND description:**

TastyFish is a company that catches and sells fish for consumption. Due to laws and regulations for commercial fishing, fishers are required to measure various lengths of each fish they catch. If the dimensions fall below a certain threshold, the fish must be released to ensure the younger population can breed. Due to the unstable motion of fishing boats, fishers are unable to weigh each fish individually because the sea’s waves prevent accurate weight measurements.

The company’s current method of weighing a fish is letting it dry and then placing it on a scale. This proves to be inefficient because the shelf life of the fish is lowered when waiting for it to dry and also during the weighing process.

In order to improve the shelf life and deliver fish at a faster rate to customers, Tasty Fish has decided to implement a machine learning solution to predict the weight of each fish given the length dimensions.

1. **Project purpose/goals:**

The purpose of the project is to develop a machine learning algorithm that will be able to accurately predict the weight of a fish using only its length dimensions. This will be achieved by training a machine learning model with a data set that contains the weight and lengths of various fish species.

1. **Descriptive method:**

A histogram will be generated from the data in order to represent the frequency of distribution for various measurements of each fish. A scatterplot will be generated to explore the correlations between different measurements and the relationship they may have to the species. A confusion matrix will be generated to assess the performance of our algorithm and the viability of its results.

1. **Predictive/Prescriptive method:**

A machine learning algorithm will be developed in order to train a model to predict the weight of a fish using supervised learning. Linear regression will be used as the regression algorithm to predict the numerical value of a fish’s weight.

**DESIGN and DEVELOPMENT:**

1. **Computer science application type (select one):**

Web app

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

Python

1. **Operating System(s)/Platform(s) you will use:**

Windows 10

1. **Database Management System you will use:**

N/A

1. **Estimated number of hours for the following:**
   * 1. Planning and Design: 12
     2. Development: 16
     3. Documentation: 14
     4. Total: 42
2. **Projected completion date:**

Estimated Date = 4/14/2023

**IMPLEMENTATION and EVALUATION:**

1. **Describe how you will approach the execution of your project:**

Python, the scikit-learn library, and Jupyter Notebook will be studied.

Linear regression algorithms will be studied.

The dataset will be downloaded and stored.

The data will be cleaned and preprocessed.

A machine learning model that uses linear regression will be developed.

Supervised learning will be used to train the model.

The model will be tested with dataset entries to ensure its accuracy.

If the model is inaccurate, we will repeat the aforementioned steps starting from the data cleaning and preprocessing.

The writeups and documentation will be created to aid in the explanation of the machine learning model.

Visualizations will be developed.

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

**STUDENT SIGNATURE**

**­­­­­­­­­­­­­­­­­­­­­­­­­­­­­** Icon

Description automatically generated

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

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

**COURSE INSTRUCTOR’S NAME:**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**COURSE INSTRUCTOR APPROVAL DATE:**