Herald College, Kathmandu

5CS037 - Concepts and Technologies of AI

`Proposal for Final Portfolio Project.`

Siman Giri {Module Leader}

Jan 12, 2025

Contents

1 Submission Guidelines

1. Instructions on Proposal Submission:

- This submission serves as a proposal for your Final Portfolio Project, outlining your intended work and approach.
- While no strict deadline is imposed for this proposal, approval by your instructor is mandatory.
- Submissions for the Final Portfolio will not be accepted unless the proposal has been reviewed and formally approved.

2. Instructions on Dataset Selection:

- Please note that datasets utilized during class exercises or pre-loaded within the scikit-learn library are not permissible for use in your project.
- You are encouraged to select a novel and appropriate dataset that aligns with the project requirements and demonstrates originality in your work.

3. Formatting Guidelines:

- Length: One page (approximately 300-400 words).
- Font: Times New Roman, 12 pt.
- Margins: 1-inch on all sides.
- Line Spacing: Single.

4. Evaluating Criteria:

Your proposal will be evaluated based on:

- Clarity and feasibility of the research question.
- Relevance and appropriateness of the dataset.
- Clear connection to a Sustainable Development Goal.
- Adherence to formatting and submission guidelines.

5. Submission:

- Submit your proposal as a PDF file. {File Name Format: `Full Name_ Uni ID.pdf`}
- Submit to your respected instructor and verify.
- There are no deadlines for this submission. However, without verification by your instructor and module leader, your final portfolio will not be accepted.

2 Guidelines for the Proposal

1. Objective of this Proposal:

Submit a one-page proposal outlining your planned final project for the course. Your project will focus on both regression and classification task and must incorporate a dataset relevant to a specific Sustainable Development Goal (SDG).

2.1 Proposal Requirements for Regression Task:

1. Research Question:

Clearly define the research question you aim to answer using regression analysis. Examples include predicting economic indicators, estimating environmental metrics, or forecasting health-related outcomes.

- **Example:** `"How can we predict air pollution levels in urban areas using meteorological and traffic data?"`

2. Dataset Description:

Provide a brief description of the dataset you plan to use. Include:

- The source of the dataset (e.g., Kaggle, government database, open data platform).
- Key features (independent variables) and the target variable.
- The size of the dataset (e.g., number of rows and columns).
- Any preprocessing or cleaning challenges you foresee.

Example:

- `"The dataset, sourced from the Global Air Quality Monitoring Network, contains over 50,000 records of hourly air pollution readings across 10 major cities. Key features include temperature, humidity, traffic volume, and industrial emissions, with PM2.5 concentration as the target variable."`

3. Connection to a Sustainable Development Goal (SDG)

Identify which SDG your project aligns with and explain the relevance of your analysis to this goal. SDGs can be found at United Nations SDGs Website. Example goals include:

- SDG 3: Good Health and Well-Being
- SDG 11: Sustainable Cities and Communities
- SDG 13: Climate Action

Example:

- `"This project aligns with SDG 11: Sustainable Cities and Communities, as it aims to provide insights into urban air quality management and support policymakers in creating healthier living environments."`

2.2 Proposal Requirements for Classification Task:

1. Research Question:

Clearly define the research question you aim to answer using classification analysis. Examples include categorizing health conditions, predicting customer behavior, or detecting fraudulent transactions.

- **Example:** `"How can machine learning models classify regions at high risk of water scarc using environmental and socioeconomic data?"`

2. Dataset Description:

Provide a brief description of the dataset you plan to use. Include:

- The source of the dataset (e.g., Kaggle, government database, open data platform).
- Key features (independent variables) and the target variable.
- The size of the dataset (e.g., number of rows and columns).
- Any preprocessing or cleaning challenges you foresee.

Example:

- `"The dataset, sourced from the United Nations Water Data Portal, contains over 20,000 records on regional water availability, population density, climate patterns, and land usage. The target variable classifies regions as 'low risk,' 'moderate risk,' or 'high risk' of water scarcity. Key features include average rainfall, agricultural land coverage, and groundwater levels. Preprocessing challenges include handling missing rainfall data for certain regions."`

3. Connection to a Sustainable Development Goal (SDG)

Identify which SDG your project aligns with and explain the relevance of your analysis to this goal. SDGs can be found at United Nations SDGs Website. Example goals include:

- SDG 6: Clean Water and Sanitation
- SDG 11: Sustainable Cities and Communities
- SDG 13: Climate Action

Example:

- `"This project aligns with SDG 6: Clean Water and Sanitation, as it focuses on identif high-risk regions for water scarcity. The insights derived from this analysis could guide policymakers in resource allocation and proactive water management strategies to ensure access to clean and sustainable water sources."`

THE - END