INFO 212: Data Science Programming I

Please be aware that this document serves as an illustrative example for the final project proposal submission. Its purpose is to provide guidance and inspiration for structuring your own proposal. We encourage you to create a customized version that aligns with your chosen dataset, research questions, and methodology.

Best wishes for your project!

Data Science Project Proposal: Predictive Analysis of Customer Behavior Using Mobile Data

1. Introduction

In this project, we aim to analyze consumer behavior using a real-world dataset derived from mobile usage. By applying data science techniques, we will uncover insights that can inform business strategies and drive decision-making.

2. Dataset Selection

We propose using the "Mobile Usage Analytics" dataset, which contains anonymized records of mobile app usage, call logs, and location data. This dataset provides a rich source of information for understanding user behavior.

3. Research Questions

Our project will address the following research questions:

1. What are the key factors influencing user engagement with mobile apps?
2. Can we predict user churn based on usage patterns?
3. How does location impact app preferences?

4. Methodology and Tools

We will employ the following methods and tools:

- **Exploratory Data Analysis (EDA)**: Investigate data distributions, missing values, and outliers.
- **Statistical Analysis**: Conduct hypothesis testing and correlation analysis.
- **Predictive Modeling**: Build machine learning models (e.g., logistic regression, decision trees) to predict user churn.

- **Visualization**: Create interactive visualizations using Python libraries (e.g., Matplotlib, Seaborn).

5. Expected Deliverables

1. Comprehensive Report:
   - Executive summary
   - Data description
   - EDA results
   - Model performance metrics
   - Insights and recommendations

2. Presentation Slides:
   - Introduction
   - Research questions
   - Methodology
   - Key findings
   - Conclusion

6. Project Timeline

1. **Data Collection and Preprocessing** (3 days): Obtain the dataset, clean missing values, and perform feature engineering.
2. **Exploratory Data Analysis** (5 days): Explore data distributions, identify patterns, and visualize insights.
3. **Predictive Modeling** (4 days): Train and evaluate machine learning models for churn prediction.
4. **Report Writing** (3 days): Document findings, methodology, and conclusions.

7. Budget

NA – Not Applicable

8. Approval and Sign-Off

Both parties (project team and stakeholders) will review and approve this proposal. Sign-off will indicate agreement on project scope, timeline, and deliverables.

Some references:

- *https://crunchingthedata.com/data-science-project-proposals/*
- *https://www.slideteam.net/blog/top-10-data-science-proposal-templates-with-examples-and-samples*