**Project Proposal: Analyzing Donor Behavior Using Deep Learning**

Project 4

Predicting Donor Behavior

Project Proposal

Caleb Cuterer

Shannon Hoffman

Krishna Pandari

Jim Pieper

Walgama Rohanajayasekara

**1. Introduction**

The purpose of this project is to analyze and predict donor behavior using machine learning techniques, specifically deep learning. By understanding patterns in donor activity, we aim to help nonprofit organizations optimize their fundraising strategies and improve donor engagement. The focus will be on identifying key factors that influence donor retention, donation frequency, and donation amounts.

**2. Data and Field of Interest**

* **Field**: Nonprofit and Charity Organizations – higher education
* **Data**: The project will utilize a dataset containing donor information, donation history, demographic data, and engagement metrics. The data may include variables such as donation amounts, donation frequency, donor age, location, and communication history.
* **Potential Data Sources**:
  + Kaggle datasets related to donor behavior and charity donations
  + Publicly available datasets from nonprofit organizations or research institutions
  + Data provided by a partnering nonprofit organization

**3. Research Questions**

The project will focus on the following key questions:

* What are the main factors that influence donor retention?
* Can we predict which donors are likely to make repeated donations?
* How can we segment donors based on their behavior and preferences?
* What patterns exist in donation amounts over time?
* How can we identify potential major donors from the existing donor base?

**4. Methodology**

* **Data Preprocessing**: Clean the dataset by handling missing values, encoding categorical variables, and normalizing numerical features.
* **Exploratory Data Analysis (EDA)**: Perform EDA to identify trends, correlations, and outliers in the data. Visualizations will be used to illustrate key findings.
* **Model Development**: Implement a deep learning model using TensorFlow/Keras to predict donor behavior. The model may include:
  + **Input Layer**: All relevant features
  + **Hidden Layers**: Multiple layers with ReLU activation functions
  + **Output Layer**: A sigmoid or softmax layer for binary/multiclass classification
* **Model Evaluation**: Evaluate the model using accuracy, precision, recall, F1 score, and AUC-ROC. Cross-validation and hyperparameter tuning will be conducted to optimize performance.
* **Donor Segmentation**: Use clustering techniques (e.g., K-means) to segment donors based on their behavior and model predictions.

**5. Expected Outcomes**

* A predictive model that identifies donors likely to make repeat donations.
* Insights into key factors influencing donor retention and donation amounts.
* Segmentation of donors into different categories based on behavior, which can be used for targeted marketing and engagement strategies.
* Recommendations for nonprofit organizations on how to improve donor retention and increase donations.

**6. Project Timeline**

* **Days 1**: Project determination, data collection
* **Day 2**: Exploratory Data Analysis/Model Development and Training
* **Day 3**: Model Evaluation and Optimization
* **Day 4**: Final Analysis/Report Writing
* **Day 5**: Presentation

**7. Potential Challenges**

* Availability of high-quality, relevant data.
* Complexity in model training due to imbalanced classes (e.g., more one-time donors than repeat donors).
* Interpreting deep learning models for actionable insights.

**8. Conclusion**

This project aims to provide actionable insights to nonprofit organizations by leveraging deep learning techniques to analyze donor behavior. The findings from this study will help organizations better understand their donor base, optimize fundraising efforts, and ultimately increase donor retention and donation amounts.

**Here’s our approach:**

**1. What are the traits of constituents that give gifts/donations?**

**a. Data Collection:**

* Gather data on various attributes, including:
  + **Demographic Information:** Age, Gender, Marital status, Education, PrimaryEmployment, City, State, Postcode, Country, Birthdate.
  + **Giving History:** Lifetime giving, First gift amount, First gift type, First gift date, Latest gift amount, Latest gift type, Latest gift date, Greatest gift amount, Greatest gift type, Greatest gift date.
  + **Affiliations and Engagement:** Constituent type, Constituent codes, Student Organizations, Arts, Fraternal Organizations, Men's Athletics, Women's Athletics, Scholarship Recipient.

**b. Data Cleaning and Preprocessing:**

* Handle missing data, correct inconsistencies, and normalize numerical fields.
* Encode categorical variables such as Gender, Education, Marital status, and PrimaryEmployment.

**c. Feature Engineering:**

* Create new features from the given data:
  + **Recency:** Calculate the number of days since the Latest gift date.
  + **Frequency:** Count the number of gifts based on First gift date, Latest gift date, and Greatest gift date.
  + **Monetary Value:** Aggregate the total amount using Lifetime giving, First gift amount, Latest gift amount, and Greatest gift amount.
  + **Engagement Score:** Sum participation in Student Organizations, Arts, Fraternal Organizations, Men's Athletics, Women's Athletics, and Scholarship Recipient.

**d. Exploratory Data Analysis (EDA):**

* Visualize and analyze the relationship between donation behavior and attributes like Age, Gender, Education, PrimaryEmployment, Constituent type, and Lifetime giving.
* Use box plots, histograms, and correlation matrices to uncover patterns.

**e. Model Building:**

* Build a predictive model (e.g., logistic regression, decision tree) to determine the traits that are most indicative of donation behavior.
* Use features such as Age, Gender, Marital status, Education, Recency, Frequency, Monetary Value, Engagement Score, and Constituent type.

**f. Interpretation:**

* Identify the most important traits that distinguish donors from non-donors using the model’s feature importances or coefficients.

**2. What are the traits of constituents that DO NOT give gifts/donations?**

**a. Data Collection:**

* Gather data on various attributes, including:
  + **Demographic Information:** Age, Gender, Marital status, Education, PrimaryEmployment, City, State, Postcode, Country, Birthdate.
  + **Giving History:** Lifetime giving, First gift amount, First gift type, First gift date, Latest gift amount, Latest gift type, Latest gift date, Greatest gift amount, Greatest gift type, Greatest gift date.
  + **Affiliations and Engagement:** Constituent type, Constituent codes, Student Organizations, Arts, Fraternal Organizations, Men's Athletics, Women's Athletics, Scholarship Recipient.

**b. Data Cleaning and Preprocessing:**

* Handle missing data, correct inconsistencies, and normalize numerical fields.
* Encode categorical variables such as Gender, Education, Marital status, and PrimaryEmployment.

**c. Feature Engineering:**

* Create new features from the given data:
  + **Recency:** Calculate the number of days since the Latest gift date.
  + **Frequency:** Count the number of gifts based on First gift date, Latest gift date, and Greatest gift date.
  + **Monetary Value:** Aggregate the total amount using Lifetime giving, First gift amount, Latest gift amount, and Greatest gift amount.
  + **Engagement Score:** Sum participation in Student Organizations, Arts, Fraternal Organizations, Men's Athletics, Women's Athletics, and Scholarship Recipient.

**d. Exploratory Data Analysis (EDA):**

* Visualize and analyze the relationship between donation behavior and attributes like Age, Gender, Education, PrimaryEmployment, Constituent type, and Lifetime giving.
* Use box plots, histograms, and correlation matrices to uncover patterns.

**e. Model Building:**

* Build a predictive model (e.g., logistic regression, decision tree) to determine the traits that are most indicative of donation behavior.
* Use features such as Age, Gender, Marital status, Education, Recency, Frequency, Monetary Value, Engagement Score, and Constituent type.

**f. Interpretation:**

* Identify the most important traits that distinguish donors from non-donors using the model’s feature importances or coefficients.

**3. Suggested steps to make a non-donor a donor?**

**a. Segmentation:**

* Segment non-donors based on traits similar to those of donors. Traits include demographic information, engagement level, and affiliations (Constituent type, Constituent codes, Education, etc.).

**b. Personalized Outreach:**

* Use segmented data to craft personalized communication strategies:
  + For younger constituents, leverage social media campaigns.
  + For those with high engagement scores, offer special events or volunteer opportunities.

**c. Engagement Strategies:**

* Increase engagement by encouraging participation in events, student organizations, and other activities to build a relationship.
* Introduce small, low-barrier giving opportunities like micro-donations.

**d. Recognition and Feedback:**

* Recognize initial gifts with personalized thank-you messages and show the impact of their contributions.

**e. Monitoring and Iteration:**

* Track the conversion rates of non-donors and continuously improve strategies based on what works.