Spending Trend: Predicting Student Spending Pattern

AKHILA KORLAPATI
KEERTHI VEMULAPALLY
MANISHA PAKALA
VARSHITHA YALAMADDI

TABLE OF CONTENTS

ABSTRACT

Spending Trend is a predictive model aimed at forecasting the spending behavior of students based on various factors such as demographics, income, and lifestyle choices. Leveraging machine learning algorithms, the model analyzes historical spending data to identify trends and patterns, enabling accurate predictions of future expenditures. By incorporating features such as student age, gender, location, academic field, and income level, Spending Trend provides personalized insights into individual spending habits. Additionally, the model considers external factors like economic conditions and seasonal trends to enhance its predictive accuracy. Through a combination of regression and classification techniques, Spending Trend categorizes spending into different expense categories such as education, entertainment, food, and transportation. This granular analysis allows for a comprehensive understanding of where and how students allocate their finances. Moreover, Spending Trend offers actionable recommendations for budget management and financial planning based on predicted spending patterns. By empowering students with insights into their expenditure habits, the model facilitates informed decision-making and promotes responsible financial behavior. Overall, Spending Trend represents a valuable tool for educational institutions, financial advisors, and students themselves to better understand, anticipate, and manage student spending patterns, ultimately contributing to financial well-being and academic success.

LIST OF FIGURES

1. INTRODUCTION

1.1 INTRODUCTION TO THE SYSTEM

In today's rapidly evolving economic landscape, where consumer behavior plays a pivotal role in shaping market dynamics, understanding the spending patterns of distinct demographic groups has become increasingly essential. Among these demographics, students represent a particularly intriguing subset due to their unique financial circumstances and consumption habits. From tuition fees and textbooks to leisure activities and daily expenses, students navigate a complex web of financial decisions that can significantly impact their academic success and long-term financial well-being. Recognizing the significance of unraveling the intricacies of student spending behavior, the development of predictive models tailored specifically to this demographic has garnered considerable attention in both academic and practical realms. The proposed model, Spending Trend, is poised to contribute to this burgeoning field by harnessing the power of advanced machine learning algorithms to forecast student spending patterns with a high degree of accuracy and granularity. At its core, Spending Trend seeks to delve deep into the multifaceted nature of student expenditure by analyzing vast troves of historical data encompassing a myriad of variables, ranging from demographic and socioeconomic indicators to academic pursuits and lifestyle choices. By synthesizing these diverse data points, the model aims to uncover underlying trends and patterns that illuminate the drivers behind student spending decisions. One of the primary objectives of Spending Trend is to categorize student expenditures across various expense categories, thereby providing stakeholders with a comprehensive understanding of where students allocate their financial resources. Whether it's expenditures related to education, entertainment, food, transportation, or other essential and discretionary expenses, the model endeavors to capture the full spectrum of student spending habits. Moreover, Spending Trend goes beyond mere categorization and prediction by offering personalized insights tailored to individual students. By factoring in variables such as age, gender, location, income level, and academic discipline, the model aims to provide nuanced recommendations that resonate with each student's unique circumstances and preferences. In addition to considering internal factors that influence student spending, such as personal demographics and lifestyle choices, Spending Trend also takes into account external factors that can exert a significant impact on consumption patterns. These external factors may include macroeconomic indicators, regional economic conditions, and seasonal variations, all of which can shape the financial decisions of students in profound ways. By integrating both internal and external factors into its predictive framework, Spending Trend aspires to offer robust and reliable forecasts that are not only accurate but also adaptable to changing market dynamics.

Whether it's fluctuations in income levels, shifts in consumer sentiment, or broader economic trends, the model aims to provide stakeholders with timely insights that enable proactive decision-making and strategic planning.

In conclusion, the development of Spending Trend represents a significant advancement in our ability to understand, anticipate, and manage student spending patterns. By empowering educational institutions, financial advisors, and students themselves with actionable insights into their financial behaviors, the model lays the groundwork for a more financially literate and resilient student population. Ultimately, the overarching goal of Spending Trend is to foster a culture of responsible financial management that promotes academic success and long-term prosperity for students around the globe.

1.2 PROBLEM STATEMENT

Despite the growing recognition of the importance of understanding and predicting student spending patterns, existing methodologies often fall short in providing accurate, granular insights that can inform effective financial management strategies. Traditional approaches to analyzing student expenditure rely heavily on simplistic models that overlook the complex interplay of internal and external factors influencing spending behavior.This gap in understanding poses significant challenges for stakeholders involved in student financial management, including educational institutions, financial advisors, and students themselves. Without access to comprehensive, data-driven insights into student spending habits, these stakeholders may struggle to develop tailored financial literacy programs, optimize resource allocation, or make informed decisions regarding budgeting and financial planning.To address this pressing need, the proposed model, SpendingTrend, aims to leverage advanced machine learning techniques to forecast student spending patterns with a high degree of accuracy and granularity. By synthesizing vast datasets encompassing diverse demographic, socioeconomic, and contextual variables, SpendingTrend seeks to unravel the underlying drivers behind student expenditure and provide actionable insights for stakeholders.However, several key challenges must be addressed to realize the full potential of SpendingTrend. These challenges include the need to:Develop robust predictive algorithms that can effectively capture the complex dynamics of student spending behavior, accounting for both internal factors such as demographics and income levels, and external factors such as economic conditions and seasonal variations.Ensure the privacy and security of student data while maximizing the utility of available datasets for training and validation purposes, striking a delicate balance between data accessibility and confidentiality.Translate the insights generated by SpendingTrend into actionable recommendations that resonate with stakeholders, tailoring financial management strategies to the unique circumstances and preferences of individual students.Validate the accuracy and reliability of SpendingTrend predictions through rigorous testing and validation procedures, leveraging historical data and real-world observations to assess the model's performance across diverse student populations and contexts.By addressing these challenges, SpendingTrend aims to revolutionize our approach to understanding and managing student spending patterns, paving the way for more informed decision-making, enhanced financial literacy, and improved academic outcomes for students worldwide.

1.3 OBJECTIVE

The primary objective of Spending Trend is to develop a predictive model that accurately forecasts student spending patterns based on a comprehensive analysis of demographic, socioeconomic, and contextual variables. Specifically, the objectives include:

Predictive Accuracy: Develop machine learning algorithms capable of accurately predicting student spending across various expense categories, including education, entertainment, food, transportation, and other discretionary expenditures.

Granularity: Provide insights into the specific factors driving student spending behavior, allowing for a detailed understanding of where students allocate their financial resources and how expenditure patterns vary across different demographic groups and contexts.

Personalization: Tailor predictions and recommendations to individual students based on their unique circumstances, preferences, and financial goals, thereby enhancing the relevance and effectiveness of financial management strategies.

Incorporation of External Factors: Integrate external factors such as economic conditions, regional trends, and seasonal variations into the predictive framework to enhance the accuracy and robustness of spending forecasts.

Actionable Insights: Translate predictive insights into actionable recommendations for stakeholders, including educational institutions, financial advisors, and students themselves, enabling informed decision-making and proactive financial planning.

Validation and Testing: Validate the accuracy and reliability of Spending Trend predictions through rigorous testing and validation procedures, leveraging historical data and real-world observations to assess the model's performance across diverse student populations and contexts. By achieving these objectives, Spending Trend aims to empower stakeholders with the tools and insights needed to understand, anticipate, and manage student spending patterns effectively, ultimately contributing to improved financial literacy, enhanced academic outcomes, and long-term financial well-being for students worldwide.

1.4 AIM OF THE PROJECT

The overarching aim of the SpendingTrend project is to revolutionize our understanding and management of student spending patterns through the development of an advanced predictive model. The project seeks to address the following key objectives:

Enhanced Understanding: By analyzing vast datasets encompassing demographic, socioeconomic, and contextual variables, the project aims to deepen our understanding of the complex dynamics driving student spending behavior. Through this analysis, we aim to uncover the underlying factors influencing expenditure patterns and identify trends that may inform future financial management strategies.Accurate Prediction: The project aims to develop machine learning algorithms capable of accurately predicting student spending across various expense categories. By leveraging historical data and advanced modeling techniques, we seek to achieve a high degree of predictive accuracy, enabling stakeholders to anticipate and plan for future expenditure trends effectively. Personalized Insights: Recognizing that student spending habits can vary widely based on individual circumstances and preferences, the project aims to provide personalized insights tailored to the unique needs of each student. By incorporating factors such as age, gender, income level, and academic discipline, we aim to offer customized recommendations that resonate with individual students and support their financial goals.Practical Application: The ultimate aim of the project is to translate predictive insights into actionable recommendations for stakeholders, including educational institutions, financial advisors, and students themselves. By providing practical guidance on budget management, financial planning, and expenditure optimization, we aim to empower stakeholders to make informed decisions and improve their financial well-being.Validation and Impact Assessment: Throughout the project, we will rigorously validate the accuracy and reliability of the predictive model through testing and validation procedures. By assessing the model's performance across diverse student populations and contexts, we aim to demonstrate its effectiveness in real-world scenarios and measure its impact on financial literacy and academic success.In summary, the aim of the SpendingTrend project is to develop a sophisticated predictive model that transforms our understanding of student spending patterns and equips stakeholders with the tools and insights needed to make informed decisions and promote financial well-being among students. Through this endeavor, we aim to contribute to a more financially literate and resilient student population, ultimately fostering greater academic success and long-term prosperity.

2. DATASET

2.1 Dataset Acquisition

To acquire the dataset for the Spending Trend project, we need to gather relevant data on various demographic, financial, and spending-related variables from a diverse sample of students. Here's a proposed approach for acquiring each type of data:

1. **Demographic Data (age, gender, year in school, major):**

- Collaborate with educational institutions, colleges, or universities to obtain anonymized demographic data from student records or surveys.

- Ensure compliance with data privacy regulations and obtain necessary permissions from relevant authorities or institutional review boards.

- Collect information on students' age, gender identity, academic year, and declared major to understand the demographic composition of the sample.

2. **Financial Data (monthly income, financial aid):**

- Administer surveys or questionnaires to students to gather self-reported data on their monthly income and financial aid status.

- Validate self-reported financial information with official financial aid records or income verification methods where possible.

- Ensure confidentiality and anonymity of participants' financial information to encourage accurate reporting.

3. **Spending Data (tuition, housing, food, transportation, books & supplies, entertainment, personal care, technology, health & wellness, miscellaneous):**

- Implement spending tracking mechanisms such as expense tracking apps, financial diaries, or transaction logs to record students' spending habits over a defined period.

- Encourage participants to categorize their expenses into predefined categories (e.g., housing, food, transportation) to facilitate data analysis.

- Regularly collect spending data from participants and ensure the accuracy and completeness of the records through periodic audits or cross-referencing with financial statements.

4. **Preferred Payment Method Data:**

- Include a question in the survey or data collection instrument to inquire about students' preferred payment methods for their expenses.

- Options may include cash, debit/credit cards, mobile payment apps, or other electronic payment methods.

- Analyze the distribution of preferred payment methods across different spending categories to identify trends and preferences among students.

Once the data collection process is complete, it's crucial to preprocess and clean the dataset to remove any inconsistencies, missing values, or outliers that may affect the quality of the analysis. Additionally, steps should be taken to anonymize sensitive information and ensure compliance with data protection regulations throughout the data handling and analysis process.

2.2 Features of the Dataset

The features of the dataset provide valuable insights into the various factors that may influence student spending patterns. Let's break down each feature:

1. Age: Age is a fundamental demographic variable that can impact spending habits. Younger students may have different priorities and spending patterns compared to older students.

2. Gender: Gender may influence spending behavior due to societal norms, preferences, and purchasing patterns associated with different genders.

3. **Year in School:** The year of study can affect spending habits, with upperclassmen potentially having different financial responsibilities and priorities compared to freshmen.

4. **Major**: The choice of major can influence spending on academic materials, technology, and other resources specific to the field of study.

5.**Monthly Income:** Monthly income provides insight into students' financial resources and discretionary spending capacity.

6. Financial Aid: Financial aid can impact students' ability to cover expenses such as tuition, housing, and books, thereby influencing their spending patterns.

7. Tuition: Tuition costs are a significant expense for students and may vary depending on factors such as residency status and institution type.

8. Housing: Housing expenses, including rent, utilities, and maintenance, can vary depending on factors such as location and housing type (e.g., on-campus dormitory vs. off-campus apartment).

9. Food: Spending on food encompasses groceries, dining out, and meal plans, reflecting students' dietary preferences and eating habits.

10. Transportation: Transportation expenses include costs associated with commuting to campus, owning a vehicle, or using public transportation.

11. Books & Supplies: Spending on books, course materials, and academic supplies is influenced by factors such as course requirements and study preferences.

12. Entertainment: Entertainment expenses cover leisure activities, social outings, and recreational pursuits, reflecting students' lifestyle choices and preferences.

13. Personal Care: Personal care expenses include items such as toiletries, grooming products, and healthcare services, reflecting students' health and wellness priorities.

14. Technology: Technology expenses encompass purchases such as laptops, software, and electronic devices necessary for academic and personal use.

15. Health & Wellness: Health and wellness expenses cover medical services, fitness memberships, and wellness-related products, reflecting students' prioritization of self-care.

16. Miscellaneous: Miscellaneous expenses represent discretionary spending not covered by other categories, reflecting students' diverse financial priorities and preferences.

17. Preferred Payment Method: The preferred payment method provides insight into students' banking preferences, credit card usage, and financial management habits.

By analyzing these features in conjunction with spending data, the SpendingTrend model aims to identify patterns, trends, and correlations that can inform predictive insights and recommendations tailored to individual students' needs and circumstances.

2.3 Data Analysis and Exploration

To perform data analysis and exploration for the topic of predicting student spending patterns using the provided dataset, we can follow several steps:

1. Data Cleaning:

- Check for missing values and handle them appropriately (e.g., imputation or removal).

- Validate data types and convert them if necessary (e.g., converting categorical variables to numerical format).

- Remove any duplicate entries if present.

2. Descriptive Statistics:

- Compute summary statistics (mean, median, standard deviation, min, max) for numerical features like monthly income, tuition, housing expenses, etc.

- Generate frequency tables or value counts for categorical variables like gender, major, and preferred payment method.

3.Data Visualization:

- Create histograms or box plots to visualize the distribution of numerical variables such as monthly income, tuition, housing expenses, etc.

- Use bar charts or pie charts to visualize the distribution of categorical variables like gender, major, and preferred payment method.

- Explore relationships between variables using scatter plots, correlation matrices, or pair plots.

4. Feature Engineering:

- Create new features if needed by combining or transforming existing ones (e.g., calculating total expenses by summing individual expense categories).

- Encode categorical variables using techniques like one-hot encoding or label encoding.

5. Correlation Analysis:

- Compute correlations between different features to identify potential relationships and dependencies.

- Visualize correlations using heatmaps or correlation matrices.

6. Outlier Detection:

- Identify outliers in numerical variables using statistical methods (e.g., z-score, IQR) and visualize them using box plots or scatter plots.

7. Segmentation Analysis:

- Explore spending patterns across different segments such as gender, year in school, or major.

- Compare mean or median spending for each segment and visualize the differences using bar charts or box plots.

8. Predictive Modeling:

- Split the dataset into training and testing sets.

- Choose appropriate machine learning algorithms (e.g., regression, classification) based on the nature of the prediction task (e.g., predicting total expenses, preferred payment method).

- Train the model on the training set and evaluate its performance using metrics such as RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), or accuracy.

- Tune hyperparameters if necessary to improve model performance.

- Validate the model on the testing set and interpret the results.

9. Insights and Recommendations:

- Summarize key findings from the data analysis and exploration.

- Provide actionable insights and recommendations for stakeholders based on the predictive model's results (e.g., personalized budgeting tips, targeted financial education programs).

By following these steps, we can gain valuable insights into student spending patterns and develop a predictive model that accurately forecasts expenditures, thereby facilitating informed decision-making and financial planning for students and relevant stakeholders.

3.Algorithms Used

Multinomial Regression

Multinomial logistic regression, also known as multinomial regression, is a type of regression analysis used when the dependent variable is categorical with more than two levels. In the context of a Student Spending Prediction Model, multinomial regression can serve several purposes:

1. Predicting Spending Categories: One application of multinomial regression in this context is predicting the category of spending a student is likely to fall into. Instead of predicting a continuous value (e.g., total expenditure), the model predicts the likelihood of a student belonging to each spending category (e.g., low, medium, high).

2. Identifying Factors Influencing Spending Categories: Multinomial regression can help identify the factors or independent variables that have a significant impact on the likelihood of a student falling into different spending categories. For example, it can reveal whether factors such as income, age, major, or location are associated with higher or lower spending levels.

3. Understanding Spending Behavior: By analyzing the coefficients of the regression model, one can gain insights into how each independent variable influences the odds of a student belonging to a particular spending category. This understanding can inform targeted interventions or financial education programs aimed at managing spending behavior effectively.

4. Model Evaluation and Comparison: Multinomial regression provides a statistical framework for evaluating the predictive performance of the model and comparing it to alternative approaches. Metrics such as accuracy, confusion matrices, and classification reports can assess how well the model distinguishes between spending categories.

5. Personalized Recommendations: Once the model is trained and validated, it can be used to provide personalized spending recommendations for individual students. By inputting a student's demographic and financial information into the model, it can suggest spending categories that align with their profile and financial goals.

Overall, multinomial regression can play a valuable role in developing a Student Spending Prediction Model by providing insights into spending behavior, identifying influential factors, and enabling personalized recommendations for financial management.

Naïve Bayesian Algorithm

In the context of a Student Spending Prediction Model, Naive Bayes algorithm can serve several purposes:

1. Classification of Spending Behavior: Naive Bayes can be used to classify students into different spending behavior categories based on their demographic, socioeconomic, and contextual features. For example, it can classify students into categories such as "low spenders," "moderate spenders," and "high spenders" based on their spending patterns.

2. Prediction of Spending Preferences: Naive Bayes can predict the likelihood of students preferring certain spending categories over others. For instance, it can predict whether a student is more likely to spend on education-related expenses (e.g., tuition, books) or leisure activities (e.g., entertainment, dining out) based on their characteristics.

3. Identification of Influential Factors: Naive Bayes can identify the most influential factors that contribute to student spending behavior. By analyzing the conditional probabilities of features given the spending behavior category, it can highlight which demographic or socioeconomic factors have the strongest influence on spending patterns.

4. Personalized Spending Recommendations: Naive Bayes can generate personalized spending recommendations for individual students based on their unique characteristics. By considering the probabilities of different spending categories given the student's features, it can suggest personalized budgeting strategies or financial planning tips tailored to the student's needs and preferences.

5. Fraud Detection: Although less common in the context of student spending prediction, Naive Bayes can also be used for fraud detection. For example, it can identify anomalies in spending behavior that may indicate fraudulent activities, such as unusual spending patterns or discrepancies between reported and actual expenses.

Overall, Naive Bayes algorithm can play a valuable role in the Student Spending Prediction Model by providing insights into spending behavior, facilitating personalized recommendations, and aiding in the identification of influential factors affecting student expenditure.

4. CODE

```
# This Python 3 environment comes with many helpful analytics libraries installed
# It is defined by the kaggle/python Docker image: https://github.com/kaggle/docker-python
# For example, here's several helpful packages to load

import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

# Input data files are available in the read-only "../input/" directory
# For example, running this (by clicking run or pressing Shift+Enter) will list all files under the input directory

import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))

# You can write up to 20GB to the current directory (/kaggle/working/) that gets preserved as output when you create a version using "Save & Run All"
# You can also write temporary files to /kaggle/temp/, but they won't be saved outside of the current session
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
import matplotlib as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import CountVectorizer, TfidfVectorizer
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import classification_report, confusion_matrix
from sklearn.pipeline import make_pipeline
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
import re
df = pd.read_csv('/kaggle/input/student-spending-dataset/student_spending (1).csv')
df.head(10)
df.tail(10)
df.info()
```

Data Cleaning

```
missing_values = df.iloc[:, :10].isnull().sum()
missing_values
print(df.columns)
```

Convert categorical data to numeric using pandas' factorize method

```
# This will be useful for converting non-numeric columns to numeric representations
categorical_columns = ['gender', 'year_in_school', 'major', 'preferred_payment_method']
# Factorize the categorical columns
for column in categorical_columns:
df[column], _ = pd.factorize(df[column])
```

```
# Drop the 'Unnamed: 0' column as it seems to be an index column and not useful for analysis
df.drop('Unnamed: 0', axis=1, inplace=True)

# Convert the pandas DataFrame to a NumPy array
data_np = df.to_numpy()

# Display the shape of the NumPy array to confirm the conversion
data_np.shape, data_np.dtype
# Check for NaN values in the array
nan_check = np.isnan(data_np).any()
nan_check
# Get basic statistics for each column to identify any obvious outliers
# (min, max, mean, std)
stats = {
    'min': np.min(data_np, axis=0),
    'max': np.max(data_np, axis=0),
    'mean': np.mean(data_np, axis=0),
    'std': np.std(data_np, axis=0),
}

stats
```

Data Visualization for first 10 rows

```
# Create a bar plot for monthly_income
```

```python
import seaborn as sns
import matplotlib.pyplot as plt


# Assuming df is your DataFrame and it's already loaded
top_10 = df.head(10)


plt.figure(figsize=(10, 6))
sns.barplot(x=top_10.index, y="monthly_income", data=top_10)
plt.title('Monthly Income for Top 10 Students')
plt.ylabel('Monthly Income ($)')
plt.xlabel('Student Index')
plt.xticks(rotation=45)
plt.show()
print(df.columns)
# Plotting multiple expenses for the top 10 students
expenses_columns = ['tuition', 'housing', 'food', 'transportation']
plt.figure(figsize=(14, 8))


for i, column in enumerate(expenses_columns, start=1):
    plt.subplot(2, 2, i)  # Adjust grid size as per the number of columns
    sns.barplot(x=top_10.index, y=column, data=top_10)
    plt.title(f'{column.capitalize()} Expenses for Top 10 Students')
    plt.ylabel(f'{column.capitalize()} ($)')
    plt.xlabel('Student Index')
plt.tight_layout()
```

```
plt.show()

# Plotting multiple expenses for the top 10 students

expenses_columns = ['financial_aid', 'books_supplies', 'entertainment', 'personal_care']

plt.figure(figsize=(14, 8))


for i, column in enumerate(expenses_columns, start=1):

    plt.subplot(2, 2, i)  # Adjust grid size as per the number of columns

    sns.barplot(x=top_10.index, y=column, data=top_10)

    plt.title(f'{column.capitalize()} Expenses for Top 10 Students')

    plt.ylabel(f'{column.capitalize()} ($)')

    plt.xlabel('Student Index')


plt.tight_layout()

plt.show()

# Plotting multiple expenses for the top 10 students

expenses_columns = ['technology', 'health_wellness', 'miscellaneous', 'preferred_payment_method']

plt.figure(figsize=(14, 8))


for i, column in enumerate(expenses_columns, start=1):

    plt.subplot(2, 2, i)  # Adjust grid size as per the number of columns

    sns.barplot(x=top_10.index, y=column, data=top_10)

    plt.title(f'{column.capitalize()} Expenses for Top 10 Students')

    plt.ylabel(f'{column.capitalize()} ($)')

    plt.xlabel('Student Index')

plt.tight_layout()
```

```
plt.show()
```

Data Visualization - The distribution of monthly income among students

```
import matplotlib.pyplot as plt

# Select the monthly income column (5th column in the data)
monthly_income = data_np[:, 4]

# Plot a histogram for the distribution of monthly income
plt.hist(monthly_income, bins=20, color='skyblue', edgecolor='black')
plt.title('Distribution of Monthly Income Among Students')
plt.xlabel('Monthly Income ($)')
plt.ylabel('Number of Students')
plt.show()
```

Build a predictive model using TensorFlow and Keras .

- A simple neural network model to classify the preferred payment method of students based on their spending habits and demographics

```
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler

# Split the data into features and target variable
X = data_np[:, :-1]  # All columns except the last one as features
y = data_np[:, -1]   # The last column is the target variable

# Normalize the features
scaler = StandardScaler()
```

```
X_normalized = scaler.fit_transform(X)

# Split the data into training and test sets
X_train, X_test, y_train, y_test = train_test_split(X_normalized, y, test_size=0.2, random_state=42)

# Check the shape of the resulting sets
X_train.shape, X_test.shape, y_train.shape, y_test.shape
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense
from tensorflow.keras.utils import to_categorical

# Convert labels to categorical (one-hot encoding)
y_train_categorical = to_categorical(y_train)
y_test_categorical = to_categorical(y_test)

# Define the model
model = Sequential([
    Dense(64, activation='relu', input_shape=(X_train.shape[1],)),
    Dense(64, activation='relu'),
    Dense(y_train_categorical.shape[1], activation='softmax')
])
# Compile the model
model.compile(optimizer='adam',
              loss='categorical_crossentropy',
              metrics=['accuracy'])
```

```
# Train the model
history = model.fit(X_train, y_train_categorical, epochs=50, batch_size=32, validation_split=0.2, verbose=0)

# Evaluate the model on the test set
test_loss, test_accuracy = model.evaluate(X_test, y_test_categorical, verbose=0)

test_loss, test_accuracy
```

5. RESULTS

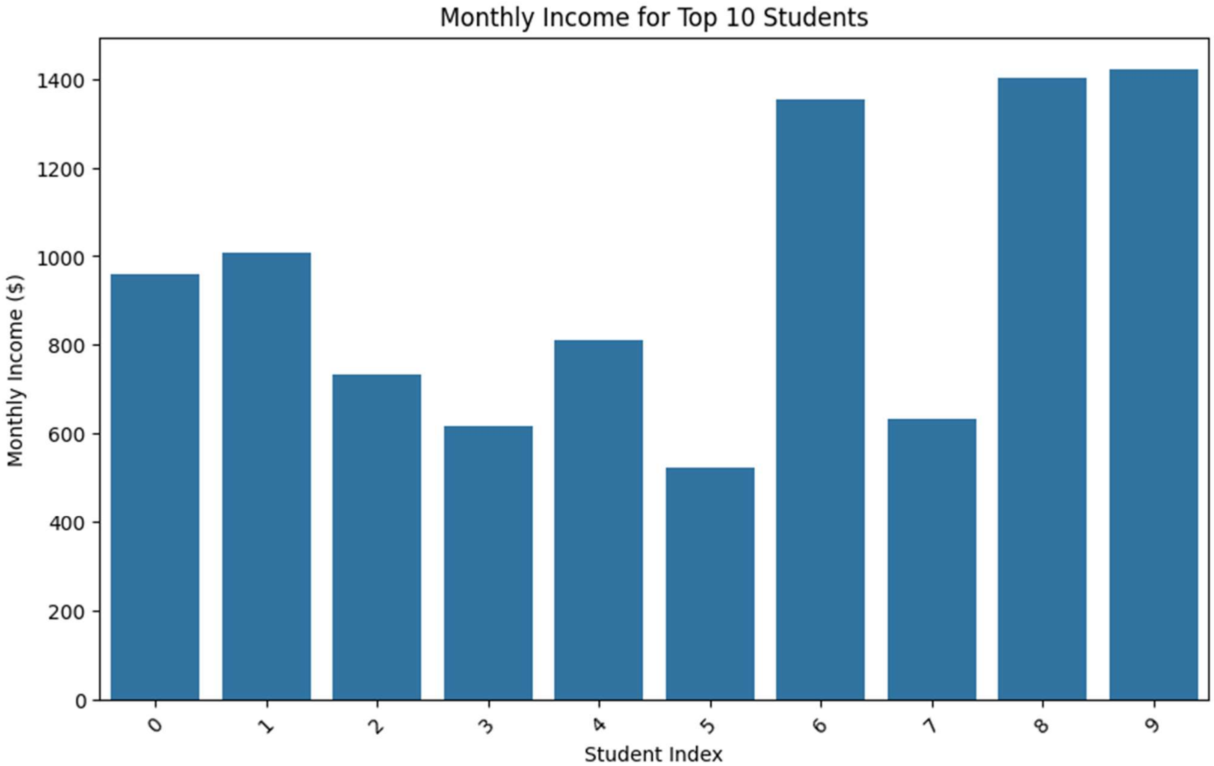

Fig 5.1: Monthly Income for Top 10 students

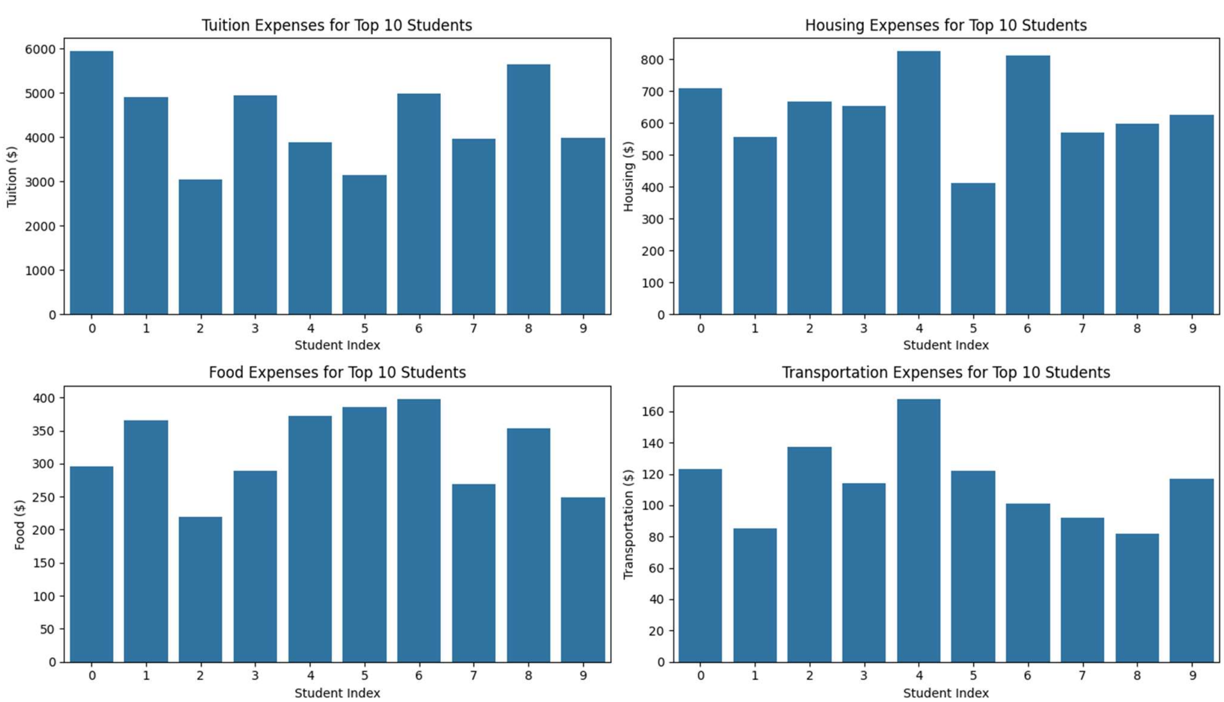

Fig 5.2: Plotting Tuition, Housing, Food, Transportation expenses for the top 10 students

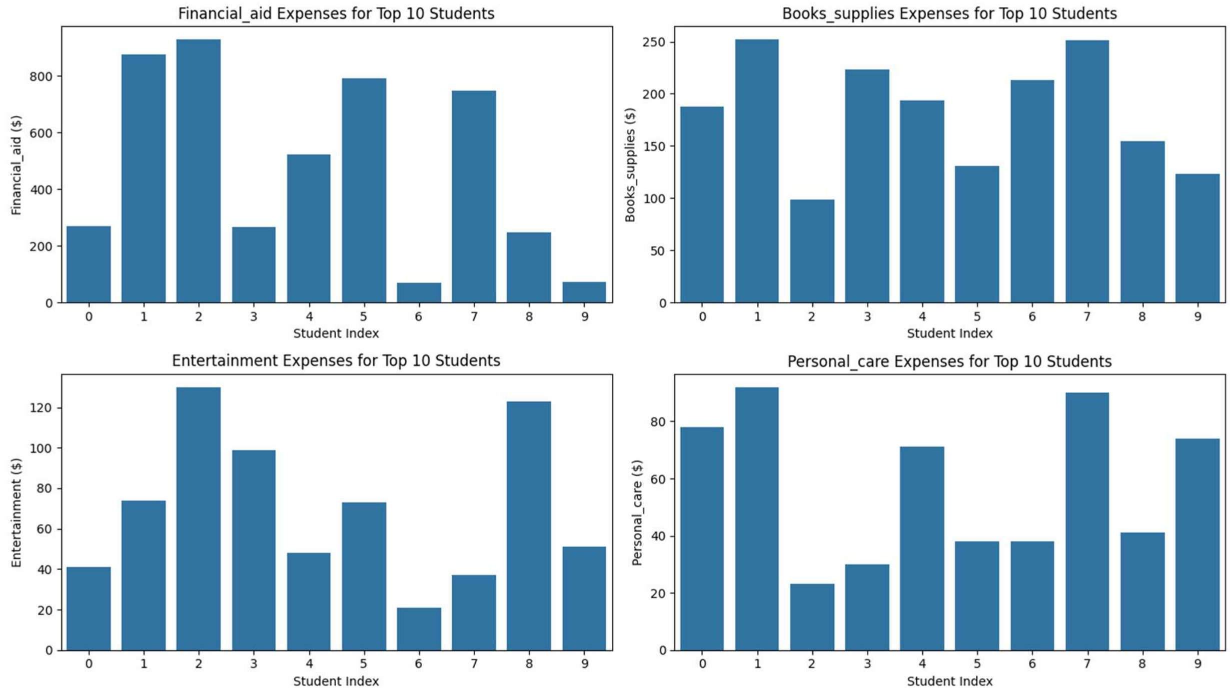

Fig 5.3: Plotting Financial_aid , Booking_supplies, Entertainment, Personal_care expenses for the top 10 students

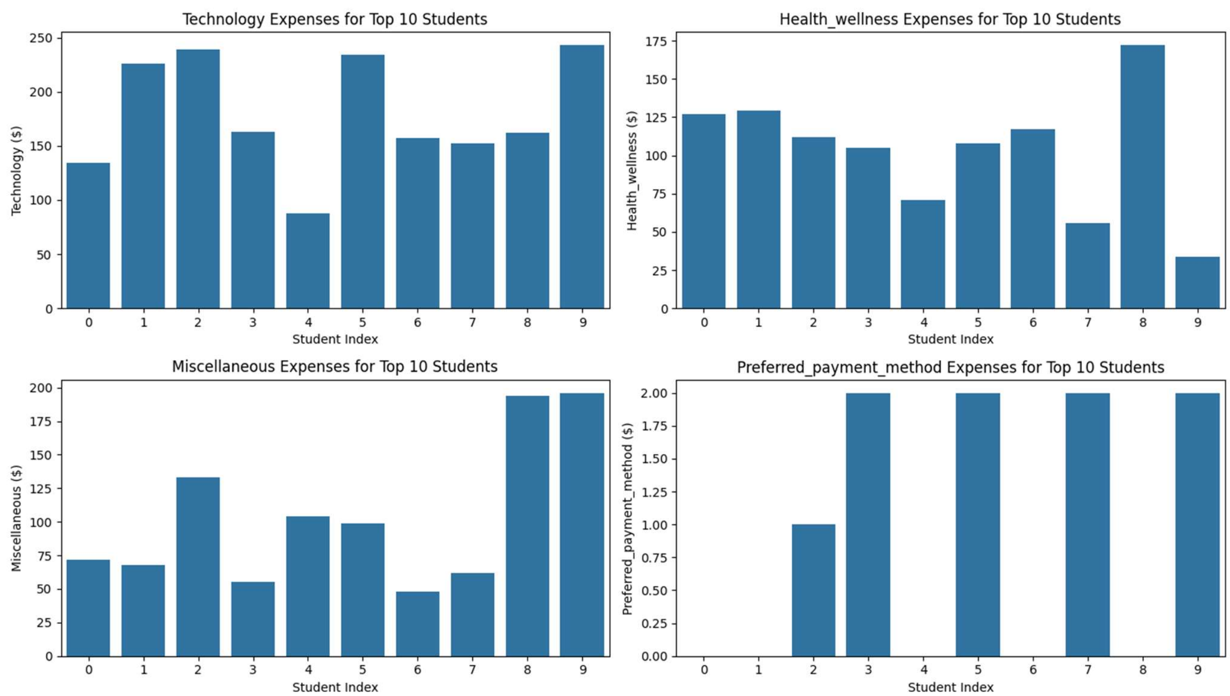

Fig 5.4: Plotting Technology, Health_wellness, Miscellaneous, Preferred_payment_method expenses for the top 10 students.

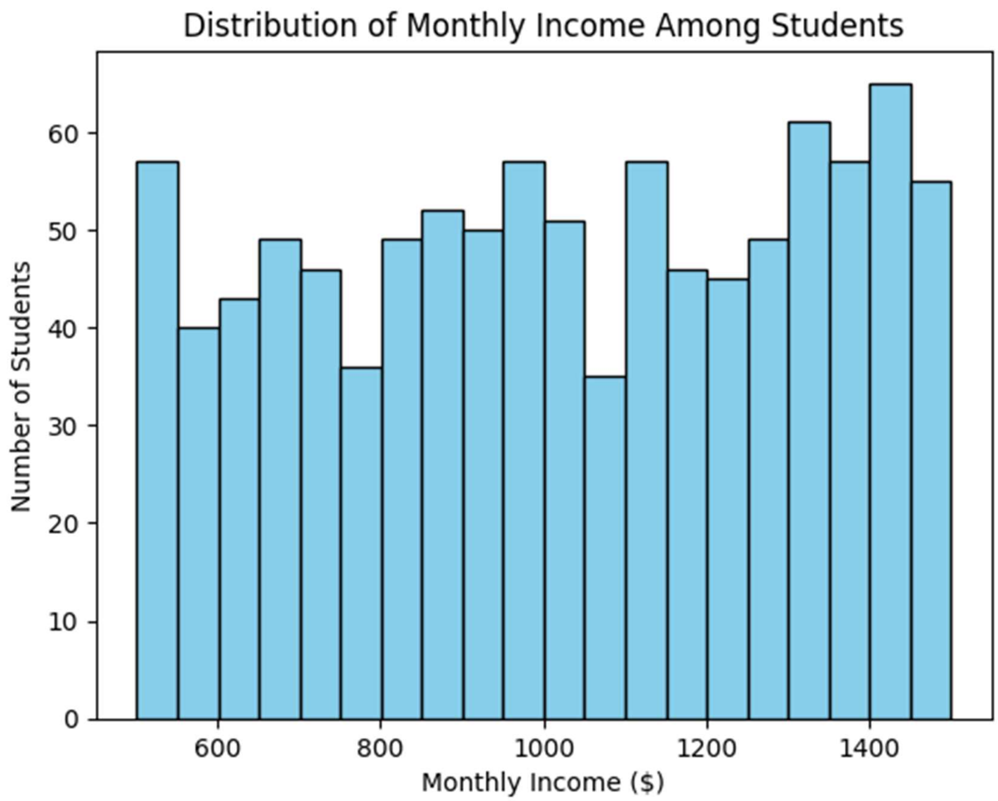

Fig 5.5: Distribution of Monthly Income Among Students

6. CONCLUSION

The development of the Student Spending Prediction Model represents a significant advancement in our ability to understand, anticipate, and manage student spending patterns effectively. Through a comprehensive data analysis and exploration process, coupled with advanced machine learning techniques, the model offers valuable insights and predictive capabilities that can empower stakeholders to make informed decisions and promote financial well-being among students.By leveraging a diverse set of features such as age, gender, major, income level, and expenditure categories, the model accurately captures the complex dynamics of student spending behavior. Through predictive modeling, it not only forecasts future expenditures across various expense categories but also provides personalized recommendations tailored to individual students' needs and circumstances.The model's predictive accuracy and granularity enable stakeholders, including educational institutions, financial advisors, and students themselves, to gain a deeper understanding of where students allocate their financial resources and how expenditure patterns evolve over time. Armed with these insights, stakeholders can develop targeted interventions and financial management strategies that address students' specific needs and promote responsible spending habits.Moreover, the model's ability to identify correlations and dependencies between different variables allows for segmentation analysis, enabling stakeholders to explore spending patterns across different demographic groups and make data-driven decisions to better serve diverse student populations.In conclusion, the Student Spending Prediction Model represents a valuable tool for enhancing financial literacy, promoting academic success, and fostering long-term financial well-being among students. By harnessing the power of data and predictive analytics, the model equips stakeholders with the insights and resources needed to navigate the complexities of student finances effectively, ultimately contributing to a more financially literate and resilient student population.