Business Analyst

Winter 2023 - INSY-695-076 - Adv Topics in Information Syst

Use Case 1 - Fairness and Ethics in ML - Airline Passenger Surveys

Ethics and fairness are critically important in AI because they are necessary to ensure that AI systems are developed and deployed in a responsible and trustworthy manner. As AI becomes more integrated into our daily lives, it has the potential to impact a wide range of social, economic, and political issues, such as privacy, security, employment, and justice.

Ensuring that AI is developed and deployed ethically and fairly requires a focus on several key areas:

- **Bias and Discrimination:** AI systems can perpetuate and amplify biases and discrimination if they are not designed and trained with fairness in mind. This can result in unfair treatment and outcomes for individuals or groups based on their race, gender, religion, or other characteristics. Therefore, it is crucial to develop AI systems that are free from bias and discrimination.

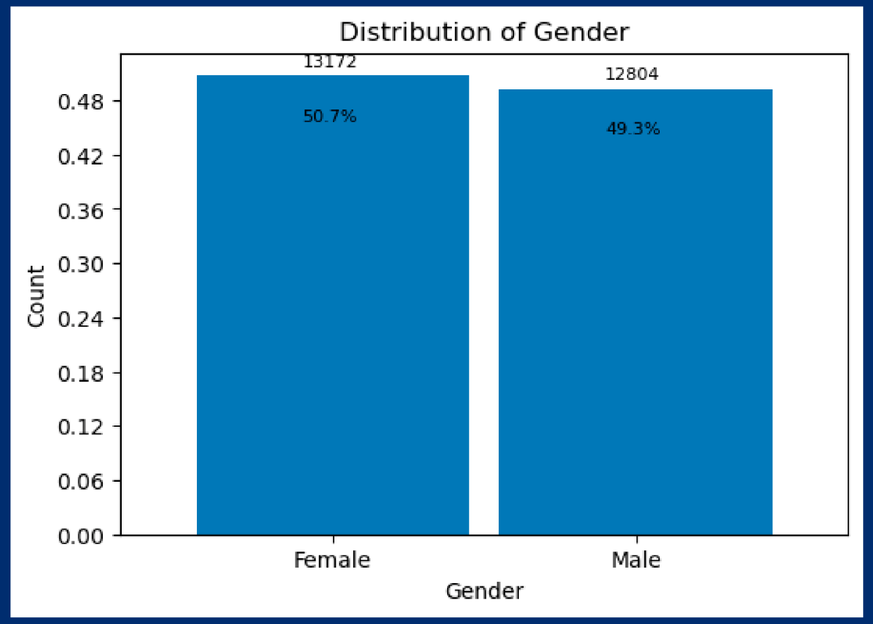
- **Transparency and Explainability:** AI systems must be transparent and explainable to build trust with users and stakeholders. This means that AI systems should be able to explain how they arrive at their decisions, what data they use, and what assumptions they make. This is important for ensuring that users can understand and evaluate the decisions made by AI systems, and for identifying and addressing any biases or errors.

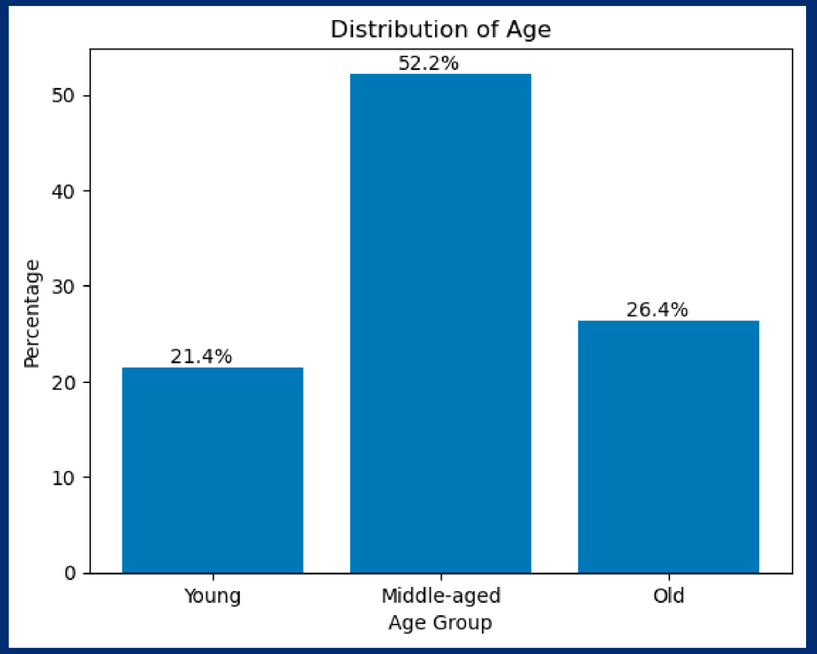
- **Privacy and Security:** AI systems can collect and process large amounts of sensitive personal data, which can raise concerns about privacy and security. Therefore, it is important to ensure that AI systems are designed and implemented with privacy and security in mind, and that they comply with relevant data protection and security standards.
- Governance and Accountability: AI systems should be subject to appropriate governance and accountability mechanisms to ensure that they are used in a responsible and ethical manner. This includes ensuring that there are clear lines of responsibility and accountability for the development, deployment, and use of AI systems, and that there are effective mechanisms for monitoring and enforcing ethical and legal standards.

- I**dentify the sensitive features in the dataset** - Sensitive features are the features that are closely associated with a particular demographic group, such as gender, race, or age. In your case, gender and age group are two potential sensitive features that you should investigate.

- **Analyze Dataset for Potential Bias** - examine the distribution of satisfaction levels across different gender or age groups to see if there are any significant differences. If so, take appropriate steps to mitig

Fairness and Ethics in ML - Airline Passenger Surveys - Continued





The results of the Exploratory Data Analysis on our dataset shows that Females are more 'neutral or dissatisfied' than males.

I**nterpretability:** Understand the reasons or issues that are disproportiantely affecting females passengers i.e female passengers are more likely to travel with women, they are morelikely to report lower levels of satisfaction.

Solution : Data Augmentation or synthetic data generation to create additional data for underrepresented age groups. This can help balance Age groups.
Fairness analysis - Define what fiarness is in the context of the model. Analyze predictions across different gender and age groups.Tracking the true positive rate and false positive rate.Biases or disparaties in model can be adjusted through the model paramters or re-sampling the data.

The respondents of the dataset are heavily cantered in the 18-50 age group with 66% of respondents being in this age group. Those who are above 50 make up approximately 25% of out dataset with the rest of respondents being under 18 .

I**nterpretability:** if model identifies age as an important feature, this may lead to Age discrimination as respondents are centered n the 18-50 category. if model recommends set assignment or meal choices based on age, it may disadvantage older passengers.

Older passengers have different needs and priorities such as seat assignment. The model may not represent their needs and concerns accurately as a result of the distribution of respondents. If model results translate into business practices this can lead to this demographic being excluded from airlines policies.

Metrics Implemented

- **Demographic Parity** - is a principle that requires that the model's predictions should not systematically favor or disfavor any demographic group.

- **True Positive Rate** - It measures the proportion of positive instances (i.e., instances belonging to the positive class) that are correctly identified by the model as positive.

- **False Positive Rate** - tells us how often the model makes a positive prediction when the true label is negative.

Metric	Age	Gender
Demographic Parity	Young: 0.25, Middle-aged: 0.47, Old: 0.45	Female: 0.41, Male: 0.43
True Positive Rate	Young: 0.85, Middle-aged: 0.94, Old: 0.93	Female : 0.925, Male : 0.932
False Positive Rate	Young : 0.00, Middle-aged : 0.00, Old: 0.00	Female: 0.00 Male: 0.00

Use Case 2 - Airline Passenger Surveys - Data/Information Leakage Analysis

Data leakage occurs when information from the test set is unintentionally included in the training set or vice versa, leading to overly optimistic performance metrics and inaccurate predictions on new data.

The implication of a Data Leakage:

- **Unforeseen precision errors of model** - the model can learn to perform well on the test set, but may not generalize well to new, unseen data. This can lead to inflated performance metrics during model evaluation, but poor performance when the model is deployed in the real world. In the context of the Airline industry, business decisions taken from the results of the model can results in financial loses, customer attrition.

- **Sensitive Information - Personal Identifiable Information (PII)** - A data leakage could result in the privacy and security of sensitive information of customers. In the case of our dataset, name and addresses are omitted, but potential leaks could include Credit card information, home addresses, cell phone numbers, race can be leaked.

- **Company Reputation** - Organizations that are found to be in violation of data privacy laws or regulations can face legal penalties and fines, and their reputation can be damaged. In addition, data leakage can erode trust in the organization or can result in regulatory fines.

Use Case 2 - Airline Passenger Surveys - Data/Information Leakage Analysis - Continued

Date Leakage: Data leakage refers to a situation where information that is not supposed to be available during training is inadvertently used to create a model. This can happen, for example, when information from the target variable is accidentally leaked into the input features, or when the training data includes information from the test data.

Data Contimination: Data contamination refers to the presence of invalid or unreliable data in a dataset, which can lead to inaccurate results and flawed conclusions. Contaminated data may result from errors during data collection, data entry, data processing or data storage. It can also be caused by the presence of outliers or anomalies, which can skew the overall distribution of the data and affect the accuracy of statistical analyses or machine learning models

Data Leakage

Data Contamination

```
No potential data leakage from Inflight wifi service into target variable
No potential data leakage from Ease of Online booking into target variable
No potential data leakage from On-board service into target variable
No potential data leakage from Cleanliness into target variable
```

```
No potential data leakage from Age into target variable
No potential data leakage from Flight Distance into target variable
Model score: 0.6505003849114703
Predicted target values:
['neutral or dissatisfied', 'satisfied', 'neutral or dissatisfied', 'satisfied',
ral or dissatisfied', 'satisfied', 'satisfied', 'neutral or dissatisfied']
Length: 5196
Categories (3, object): ['dissatisfied', 'neutral or dissatisfied', 'satisfied']
```