User Behavior Analytics for Anomaly Detection

Karan Chawla
karancha@usc.edu

Project Overview

Objective: Analyze user activity data to detect anomalous behavior using Machine Learning

Importance: Enhances security by identifying potential insider threats

Key Components:

- **Data Generation**
- **Anomaly Detection**
- **Predictive Analysis**
- **Visualization**

Data Generation

Created synthetic user activity data

- **Number of users: 100**
- **Duration: 30 days**
- **Actions: login, logout, file_access, email_sent, data_upload**

Saved data to user_activity.csv

Anomaly Detection

- Loaded user activity data
- Performed feature engineering
- Trained Isolation Forest model
  a. Identified normal and anomalous users
  b. Balanced the dataset
- Saved results to balanced_users.csv and anomalous_users.csv

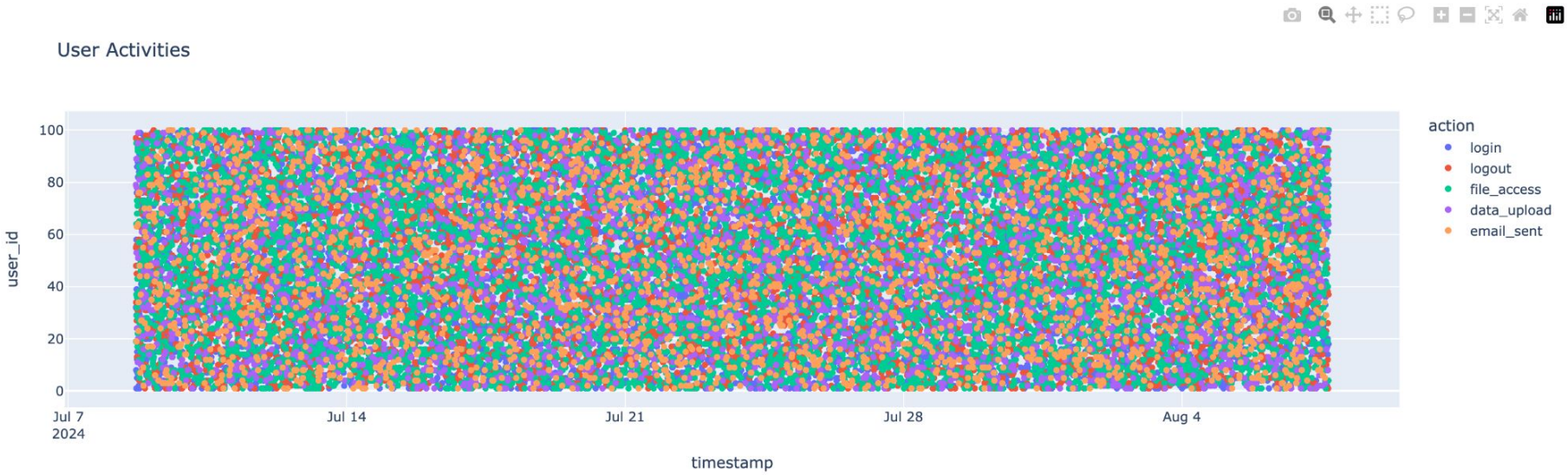
Predictive Analysis

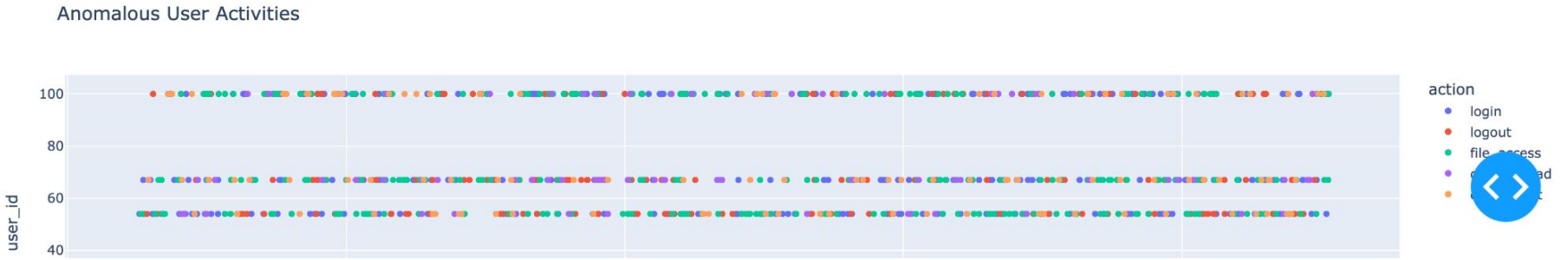
- **Loaded balanced user data**
- **Generate labels (0: normal, 1: anomalous)**
- **Trained logistic regression model**
- **Evaluated model performance**
- **Accuracy: 1.0**

Visualization Dashboard

- **Created Dash web application**
- **Visualized user activities and anomalies**
- **Scatter plots showing user actions and anomalous activities**
- **Enhanced understanding of user behavior patterns**

User Behavior Analytics Dashboard





Results and Key Features

- Successfully detected anomalous user behaviors
- Provided visual insights into user activities
- Enhanced security monitoring capabilities
- Synthetic data generation for testing
- Machine learning-based anomaly detection
- Predictive analysis for user behavior
- Interactive web-based visualization

Conclusion

- Summary of the project
  - Demonstrated machine learning for anomaly detection
  - Created a comprehensive visualization dashboard
- Future Work
  - Integrate with real-time data
  - Enhance model accuracy with additional features
- Technologies used
  - Python
  - Libraries: pandas, numpy, scikit-learn, dash, plotly

Thank you!