Week 12: AI in Data Science – Industry Case Study

Introduction:

Artificial Intelligence (AI) has become a critical tool in modern data science, especially in industries that handle large volumes of data and require fast decision-making. One of the most impactful applications of AI is in the financial sector, particularly in credit card fraud detection. Financial institutions face significant challenges due to increasing digital transactions and sophisticated fraud techniques. This project focuses on applying machine learning and deep learning techniques to detect fraudulent credit card transactions efficiently and accurately.

This case study connects the developed Credit Card Fraud Detection model with real-world industry applications, aligning with concepts discussed in AI in Finance (MIT) and global AI practices.

Problem Statement: Fraud Detection in the Financial Industry

Credit card fraud is a major concern for banks and financial institutions. Fraudulent transactions can result in:

- Financial losses for banks and customers
- Loss of customer trust
- Legal and regulatory penalties
- Increased operational costs

Traditional rule-based systems struggle to keep up with evolving fraud patterns because they rely on predefined rules. These systems often fail to detect new fraud strategies and generate high false positives. Therefore, AI-based models are required to automatically learn complex transaction patterns and identify fraud in real time.

Dataset and Data Preparation

The project uses the **Credit Card Fraud Detection dataset**, which contains anonymized transaction features and a target variable indicating whether a transaction is fraudulent (Class = 1) or non-fraudulent (Class = 0).

Key preprocessing steps included:

- Removing unnecessary columns such as TransactionID and TransactionDate
- Handling missing values and duplicate records
- Scaling numerical features using StandardScaler
- Addressing class imbalance using stratified data splitting

Proper data preparation ensured the dataset was suitable for machine learning and deep learning models.

AI Models Used in the Project

Several models were implemented throughout the project to improve fraud detection accuracy:

1. **Logistic Regression**
   Used as a baseline classification model.
2. **Random Forest Classifier**
   Improved performance by capturing non-linear relationships.
3. **Artificial Neural Networks (ANN)**
   - Baseline ANN with two hidden layers
   - Advanced ANN with three hidden layers (64, 32, 16 neurons)
   - ReLU activation and Adam optimizer

The advanced ANN demonstrated improved ability to detect complex fraud patterns compared to traditional models.

Model Evaluation and Performance

The models were evaluated using industry-relevant metrics:

- Precision
- Recall
- F1-score
- ROC-AUC score
- Confusion Matrix

Because fraud datasets are highly imbalanced, accuracy alone was not sufficient. Emphasis was placed on recall and ROC-AUC to ensure fraudulent transactions were detected effectively. The advanced ANN achieved strong performance, making it suitable for real-world deployment scenarios.

Industry Application: Real-Time Fraud Detection

In a real-world financial environment, this AI-based fraud detection system can be deployed as follows:

- Incoming credit card transactions are passed through the trained model
- The model predicts the probability of fraud in real time
- High-risk transactions are flagged for manual review or automatically blocked
- Low-risk transactions are approved instantly

This approach allows banks to prevent fraud proactively while minimizing inconvenience to legitimate customers.

Benefits to Financial Institutions

The application of AI in fraud detection provides multiple benefits:

- **Reduced financial losses** due to early fraud detection
- **Improved customer trust and security**
- **Scalable solutions** that handle millions of transactions daily
- **Lower false positives** compared to rule-based systems
- **Faster decision-making** through automation

These benefits align with industry practices discussed in AI in Finance (MIT).

Ethical Considerations and Responsible AI

Ethical AI is crucial in financial applications. This project considers:

- Transparency through explainable AI (SHAP)
- Fairness by avoiding biased decision-making
- Accountability by tracking and explaining model predictions
- Data privacy by using anonymized datasets

Responsible AI ensures trust, compliance, and ethical deployment in real-world systems.