Deep Learning Assignment 2 Report

5 Representative Applications of Supervised Learning

1. Medical Diagnosis

• Used in detecting the presence or absence of diseases using imaging or clinical measurements..
• A typical target is always a binary or multi-class label (e.g. a malignant or a benign).

2. Credit-Card Fraud Detection

• Supervised classifications models (e.g., XGBoost) are trained on millions of labelled card transaction records to predict if a transaction is fraud in real time at the payment gateway.
• They achieve >= 90% recall on the known fraud patterns. This helps in auto-blocking suspicious card transactions within milliseconds.

3. Water Quality Prediction (Regression/Classification models)

• Regression and classification models are trained on sensor data(historical) like ph, Turbidity, Heavy metals concentration, etc to predict the Water Quality Index(WQI) or binary labels of safe or unsafe.
• High Accuracy(99.9%) helps the authorities to predict the water quality and issue early warnings before contamination spreads out.

4. Sentiment Analysis on Product Reviews

• Predict star rating (1–5) or polarity (positive/negative) from raw text(text classification).
• Achieves 95% accuracy on product/movie reviews and helps brands for real time product monitoring and helps in automating customer support.

5. Forest Fire Prediction

• The model is trained on historical forest fire records with labeled data (e.g., weather conditions, temperature, humidity → “fire” or “no fire”).
• The trained model uses current environmental data to predict the likelihood of a forest fire, helping in early warning and prevention.

Forest Fire Prediction Model Performance Analysis

Dataset used: Montesinho Park Forest Fires Dataset
Total Records: 517
Features: 12(Weather conditions, spiral coordinates, temporal data)
Target Variables: Fire burned area (hectares)

Model Performance Results

Classification Models (Predicting Fire Occurrence: Yes/No)

Model	Accuracy	Performance Level
MLP Classifier	**55.80%**	**Best Overall**
Logistic Regression	53.59%	Moderate

Regression Models (Predicting Burned Area in Hectares)

Model	R^2 Score	Performance Level
MLP Regressor	0.061	Poor
Linear Regression	0.021	Very Poor
Decision Tree Regressor	-11.23	Failed
Random Forest Regressor	-588.77	Catastrophic Failure

Best model : MLP Classifier - 55.80%

Optimal Task Framing

The MLP Classifier reframed the prediction problem from continuous regression (exact burned area) to binary classification (fire occurrence). This helped in:

- Eliminating the extreme variance problem in the fire size prediction.
- Creating a more balanced target distribution (47% no fire vs. 52% fire).
- Aligning with practical fire management needs.

Architecture Advantages

- The multi-layer perceptron architecture improved pattern recognition capabilities.
- Non-linear activation functions captured intricate interactions between weather variables.
- Automatically discovered relevant weather pattern combinations.

Weather Pattern Recognition

- Being based on neural networks, it was able to capture complex interactions between humidity, temperature, and wind speed.

- Non-obvious relationships between fire weather indices were captured.

Handling Data Characteristics

- Less sensitive to outliers than the tree-based models.
- Benefited from standardized input features.
- Controlled overfitting through regularization built into the architecture design.