

1.A bank wants to predict whether a loan applicant will default based on credit score, income, and past loan history. What type of ML problem is this, and what steps would you take to solve it?

**Problem Type:** Supervised Learning – **Classification**

**Steps to Solve:**

1. Collect labeled data: credit score, income, past loan history, and whether the applicant defaulted.
2. Clean data: handle missing values, normalize/scale numeric features.
3. Encode categorical variables if any.
4. Split data into training and test sets.
5. Train a classification model (e.g., Logistic Regression, Random Forest).
6. Evaluate using accuracy, precision, recall, AUC.
7. Deploy to predict future applicants' risk.

2.A retail store wants to predict the demand for different products to optimize inventory levels. What type of ML problem is this, and what steps would you take to solve it?

**Problem Type:** Supervised Learning – **Regression** (or Time Series Forecasting)

**Steps to Solve:**

1. Collect historical sales data, seasonal trends, holidays, promotions, etc.
2. Preprocess: handle missing values, feature engineering for dates.
3. Train/test split or use time series validation (e.g., sliding window).
4. Train a regression model (e.g., Linear Regression, XGBoost, ARIMA).
5. Evaluate using MAE, RMSE.
6. Forecast future demand for inventory optimization.

3.A factory wants to detect whether a manufactured product is defective based on sensor readings and quality control data. What type of ML problem is this, and what steps would you take to solve it?

**Problem Type:** Supervised Learning – **Classification**

**Steps to Solve:**

1. Gather sensor and quality control data with labels (defective or not).
2. Clean and normalize data.
3. Split into training and test sets.
4. Train a classification model (e.g., SVM, Random Forest).
5. Evaluate using precision (important in defect detection), recall, F1-score.
6. Deploy for real-time quality control alerts.

4. A healthcare provider wants to analyze patient symptoms and classify them into different disease categories. What type of ML problem is this, and what steps would you take to solve it?

**Problem Type:** Supervised Learning – **Multi-class Classification**

**Steps to Solve:**

1. Collect patient symptom data and disease labels.
2. Preprocess: encode symptoms, clean data, impute missing values.
3. Split into training and test sets.
4. Train a multi-class classifier (e.g., Decision Tree, Neural Network).
5. Evaluate using accuracy, precision, confusion matrix.
6. Deploy in clinical decision support systems.

5. An e-commerce company wants to identify and remove fake reviews posted by bots or fraudsters. What type of ML problem is this, and what steps would you take to solve it?

**Problem Type:** Supervised Learning – **Classification**

**Steps to Solve:**

1. Collect labeled review data (real vs. fake).
2. Preprocess text (cleaning, tokenization, vectorization – e.g., TF-IDF).
3. Add metadata (e.g., user activity, account age).
4. Split into training/test sets.
5. Train a text classification model (e.g., Logistic Regression, BERT).
6. Evaluate using precision, recall, F1-score.
7. Monitor and retrain regularly due to evolving fraud tactics.

6. A financial firm wants to predict stock price movements based on historical price data and market indicators. What type of ML problem is this, and what steps would you take to solve it?

**Problem Type:** Supervised Learning – **Regression (or Time Series Prediction)**

**Steps to Solve:**

1. Collect historical price data and technical indicators (RSI, MACD, etc.).
2. Create lag features, normalize data.
3. Use time-aware train/test split.
4. Train a regression or time-series model (e.g., LSTM, ARIMA, XGBoost).
5. Evaluate using RMSE, directional accuracy.
6. Use predictions as signals for trading decisions.

7.A social media platform wants to detect fake user accounts based on user activity and profile data. What type of ML problem is this, and what steps would you take to solve it?

**Problem Type:** Supervised Learning – **Classification**

**Steps to Solve:**

1. Collect user profile and activity data with labels (real/fake).
2. Clean and preprocess: extract features like login frequency, message volume.
3. Train/test split.
4. Train a classification model (e.g., Random Forest, Gradient Boosting).
5. Evaluate using ROC-AUC, F1-score.
6. Deploy to flag suspicious accounts in real-time.

8.A marketing agency wants to segment customers into different groups based on their purchasing behavior. What type of ML problem is this, and what steps would you take to solve it?

**Problem Type:** Unsupervised Learning – **Clustering**

**Steps to Solve:**

1. Collect purchase behavior data (e.g., frequency, basket size, recency).
2. Normalize data (especially when using K-Means).
3. Apply clustering algorithm (e.g., K-Means, DBSCAN).
4. Determine optimal number of clusters (Elbow Method, Silhouette Score).
5. Interpret clusters and label them (e.g., high-value, occasional).
6. Use segments for targeted marketing.

9.A geospatial research team wants to analyze satellite images to classify different land types (forest, water, urban). What type of ML problem is this, and what steps would you take to solve it?

**Problem Type:** Supervised Learning – **Image Classification**

**Steps to Solve:**

1. Collect labeled satellite imagery with land type labels.
2. Preprocess images (resizing, augmentation).
3. Split into train/validation/test sets.
4. Train a CNN model (e.g., ResNet, EfficientNet).
5. Evaluate using accuracy, precision, confusion matrix.
6. Deploy to classify new images for mapping and analysis.

10. A streaming service wants to predict which users are likely to cancel their subscriptions. What type of ML problem is this, and what steps would you take to solve it?

**Problem Type:** Supervised Learning – **Classification**

**Steps to Solve:**

1. Collect user behavior data (e.g., usage frequency, subscription status).
2. Preprocess data, encode categorical features, handle class imbalance.
3. Split into train/test sets.
4. Train classification models (e.g., Logistic Regression, XGBoost).
5. Evaluate using precision, recall, AUC.
6. Use model to identify at-risk users and take retention actions.