## 1. Understand the Business Problem

- What is the problem?
- Why is this problem important?
- Who benefits from the solution?
- What are the success metrics?
    - Success metrics measure how effectively the machine learning solution impacts the **business goals** or **real-world outcomes**.

## 2. Translate Business Problem to ML Problem

- Is it a prediction problem? (Supervised Learning)
- Is it about uncovering patterns? (Unsupervised Learning)
- Does it involve a sequence of actions? (Reinforcement Learning)

- Identify Outputs:
    - What is the target variable? (E.g., churn: Yes/No)
    - Is it regression (continuous value) or classification (categories)?

## 3. Current solution?

## 4. Plan the Data Pipeline

- What data is needed? (E.g., demographics, purchase history, interactions)
- Where will the data come from? (Databases, APIs, surveys, web scraping)
- Data Cleaning, feature engineering

## 5. Choose the Evaluation Metric

- Evaluation metrics are methods or standards used to measure how well a machine learning model is performing.

Common Metrics for Machine Learning Problems:

### 1. Classification Metrics

Used for tasks where the target variable is a category (e.g., churn: Yes/No).

- **Accuracy:** Proportion of correct predictions.
- **Precision:** Focuses on how many predicted positives (churners) were correct.
- **Recall (Sensitivity):** Measures how many actual positives were identified.
- **F1-Score:** Balances Precision and Recall.
- **ROC-AUC:** Evaluates the model's ability to distinguish classes.

### 2. Regression Metrics

Used for predicting continuous values (e.g., revenue, temperature).

- **Mean Absolute Error (MAE):** Average of absolute errors.
- **Mean Squared Error (MSE):** Penalizes larger errors more than MAE.
- **$R^2$ (R-squared):** Proportion of variance in the target explained by the model

**7. Online vs. Batch?**

**8. Check assumption**