Notes

A/B Testing

* A/B Testing (also called split testing) is like a science experiment for decision-making.
* It helps us compare two versions of something (like a website, email, or a stock trading strategy) to see which one performs better.

Simple Example:

* Imagine you are a YouTuber, and you want to know which thumbnail gets more clicks on your video.
  + Option A 🡪 A bright, colorful thumbnail.
  + Option B 🡪 A dark, minimalistic thumbnail.
* You show half of your viewers Thumbnail A and the other half Thumbnail B. After a few days, you compare which one gets more clicks.

Steps Involved in A/B Testing:

* Decide what you want to improve (The Goal).
* Create two Versions.
* Split your audience randomly.
* Collect Data.
* Analyze the Results.
* Pick the Winner.

Where is A/B Testing used?

* It is used everywhere.
  + E-Commerce
  + Email Marketing
  + Apps & Websites
  + Stock Market Trading

Key Terms in A/B Testing

* Control Group (A): The original version (no changes).
* Treatment Group (B): The new version (with changes).
* Conversion Rate: The percentage of people who take action.
* Statistical Significance: A mathematical test to make sure the difference is real, not luck.
* P-value: A number that tells us if the results are reliable.

Example:

* A/B Testing for Stock Market Predictions
  + A quantative trading firm wants to improve its stock price prediction model using Gen AI.
  + They currently use a traditional machine learning model but believe that a Large Language Model fine-tuned on financial news might improve accuracy.
* The Hypothesis
  + Null Hypothesis (H0): The Gen AI model does not improve stock price prediction accuracy.
  + Alternative Hypothesis (H1): The Gen AI model improves stock price prediction accuracy.
* To test this, the firm compares two models using an A/B test:
  + Model A (control) 🡪 Traditional ML-Based stock price predictor.
  + Model B (Treatment) 🡪 LLM-enhanced stock price predictor.
* Generative AI Model Setup:
  + Model A: Traditional ML Approach
    - Uses historical stock prices.
    - Applies Random Forest or XGBoost.
    - Predicts next-day stock price based on past data.
  + Model B: Generative AI + Financial News
    - Uses historical stock prices.
    - Fine-tuned on financial news and earnings reports.
    - Uses an LLM to extract market sentiment.
    - Predicts next-day stock price.
* Step 1: Data Collection.
  + We test the models over 100 stocks for 30 trading days and measure prediction accuracy.
  + Each model predicts next-day closing prices, and we compare them to actual prices.
  + Metrics Tracked:
    - Mean Absolute Error (MAE) 🡪 Measures how far predictions are from actual prices.
    - Root Mean Squared Error (RMSE) 🡪 Penalizes large errors more.
    - R-Squared 🡪 Measures how well the model explains price movements.
* Step 2: Running the A/B Test.
  + We split the 100 stocks into two groups:
    - 50 stocks use Model A (ML only).
    - 50 stocks use Model B (Gen-AI enhanced).
* Step 3: Results

|  |  |  |  |
| --- | --- | --- | --- |
| Metric | Model A | Model B | Improvement? |
| MAE | 2.15 | 1.78 | 17% lower (Better) |
| RMSE | 3.05 | 2.61 | 14% lower (Better) |
| R-Squared | 0.72 | 0.81 | Higher (Better fit) |

Conclusion:

* The Generative AI Model significantly reduces prediction errors and better captures market trends.
  + Lower MAE & RMSE 🡪 More accurate predictions.
  + Higher R-Squared 🡪 Model understands stock price movement better.