**Tennis mathcesDetector**

**Requirements**

* **Match Data:** Historical match results, including players, scores, and tournament details.
* **Player Stats:** ATP/WTA rankings, head-to-head records, win/loss ratios, and recent performance.
* **Surface Information:** Clay, grass, hard court, etc., as different players perform differently on each.
* **Injury Reports:** Recent injuries or withdrawals affecting player performance.
* **Weather Conditions:** Wind speed, temperature, and humidity, if outdoor matches are included.

**Model requirements**

* **Input Features:** Player rankings, recent form, head-to-head record, surface preference, etc.
* **Target Variable:** Binary classification (win/loss) or probabilistic outcome (likelihood of winning).
* **Feature Engineering:**
* Moving averages of player performance over the last X matches.
* Elo rating system for player strength adjustments.
* Tournament importance (e.g., Grand Slam vs. ATP 250).

**Training & Evaluation**

* **Data Splitting:** Train/test split (e.g., 80/20) with a time-aware validation strategy.
* **Metrics:** Accuracy, F1-score, log loss (for probabilistic models), AUC-ROC.
* **Baseline Model:** Compare against a simple heuristic (e.g., higher-ranked player wins).

**Deployment Requirements**

* **Live Updates:** Ability to incorporate recent match results dynamically.
* **Explainability:** Feature importance insights (e.g., SHAP values).
* **Model Retraining:** Mechanism to update with fresh data periodically.