UNIVERSITAS
GADJAH MADA
**Comparative Evaluation of Machine Learning and Deep Learning Models for Major League Baseball Game
Outcome Prediction**
Taichi Mori, Drs. Edi Winarko, M.Sc., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

INTISARI

EVALUASI KOMPARATIF METODE MACHINE LEARNING DAN DEEP LEARNING UNTUK PREDIKSI HASIL PERTANDINGAN MAJOR LEAGUE BASEBALL

Oleh

TAICHI MORI
21/483407/PA/21066

Penelitian ini menyajikan kerangka kerja perbandingan machine learning untuk memprediksi hasil pertandingan Major League Baseball (MLB) menggunakan data historis terstruktur. Penelitian ini mengevaluasi beberapa model supervised learning, termasuk logistic regression, decision tree, random forest, XGBoost, support vector machine, artificial neural network, dan one-dimensional convolutional neural network, untuk memprediksi keluaran biner berupa menang atau kalah.

Beberapa dataset MLB yang tersedia secara publik, termasuk Lahman's Baseball Database, game log Retrosheet, sabermetrik FanGraphs yang diakses melalui pustaka pybaseball, data pelacakan pitch Statcast, dan data historis peluang taruhan, diintegrasikan untuk membangun dataset terpadu pada tingkat pertandingan sebanyak 26.822 observasi tim-pertandingan dari enam musim reguler, yaitu 2018 hingga 2023. Dataset ini mencakup statistik tim, indikator performa terkini, fitur sabermetrik lanjutan, dan metrik kualitas kontak Statcast.

Metodologi penelitian meliputi pengumpulan data, prapemrosesan, rekayasa fitur, implementasi model, penyetelan hyperparameter melalui validasi silang lima lipatan, serta evaluasi kinerja menggunakan accuracy, F1-score, dan Area Under the ROC Curve (AUC-ROC). Kerangka eksperimen yang terstandar diterapkan pada seluruh model untuk memastikan perbandingan yang adil dan dapat direproduksi.

Hasil penelitian menunjukkan bahwa Logistic Regression memperoleh AUC-ROC tertinggi sebesar 61,50%, sedangkan 1D-CNN memperoleh accuracy tertinggi sebesar 58,40%. Seluruh tujuh model berada pada rentang kinerja yang sempit, yaitu 54–59% untuk accuracy dan 55–62% untuk AUC-ROC. Berlawanan dengan hipotesis bahwa model yang lebih ekspresif akan mendominasi, ANN mencatat kinerja terendah secara keseluruhan (AUC = 55,75%), yang menunjukkan bahwa kompleksitas model tidak selalu menjamin hasil yang lebih baik pada data tabular terstruktur dengan jumlah fitur terbatas. Analisis feature importance mengidentifikasi fitur pasar taruhan (`moneyLine`, `projectedRuns`) sebagai prediktor yang

UNIVERSITAS
GADJAH MADA
Comparative Evaluation of Machine Learning and Deep Learning Models for Major League Baseball Game Outcome Prediction
Taichi Mori, Drs. Edi Winarko, M.Sc., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

paling dominan, diikuti oleh fitur rolling runs yang direkayasa.

Penelitian ini menunjukkan bahwa dataset terpadu dari berbagai sumber yang dikombinasikan dengan kerangka evaluasi yang konsisten dapat memberikan wawasan yang bermakna dan dapat direproduksi mengenai kinerja model prediktif dalam tugas analitik olahraga. Kerangka kerja yang diusulkan memberikan metodologi yang modular dan dapat diperluas untuk diterapkan pada olahraga tim lainnya, seperti sepak bola dan bola basket.

Kata kunci: Major League Baseball, analitik olahraga, machine learning, deep learning, prediksi hasil pertandingan, studi komparatif

UNIVERSITAS
GADJAH MADA
**Comparative Evaluation of Machine Learning and Deep Learning Models for Major League Baseball Game
Outcome Prediction**
Taichi Mori, Drs. Edi Winarko, M.Sc., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

COMPARATIVE EVALUATION OF MACHINE LEARNING AND DEEP LEARNING MODELS FOR MAJOR LEAGUE BASEBALL GAME OUTCOME PREDICTION

By

TAICHI MORI
21/483407/PA/21066

This study presents a comparative machine learning framework to predict the outcomes of Major League Baseball (MLB) games using structured historical data. The research evaluates several supervised learning models, including logistic regression, decision trees, random forests, XGBoost, support vector machines, and neural networks, to predict binary outcomes (win or loss).

Multiple publicly available MLB datasets, including Lahman's Baseball Database, Retrosheet game logs, FanGraphs sabermetrics accessed via the pybaseball library, Statcast pitch-tracking data, and historical betting odds, were integrated to construct a unified game-level dataset of 26,822 team-game observations spanning six regular seasons from 2018 to 2023. The dataset incorporates team statistics, recent form indicators, advanced sabermetric features, and Statcast contact-quality metrics.

The methodology encompasses data collection, preprocessing, feature engineering, model implementation, hyperparameter tuning via five-fold cross-validation, and performance evaluation using accuracy, F1-score, and Area Under the ROC Curve (AUC-ROC). A standardized experimental framework was applied across all models to ensure fair and reproducible comparison.

The results reveal that Logistic Regression achieved the highest AUC-ROC (61.50%), while the 1D-CNN achieved the highest accuracy (58.40%), with all seven models performing within a narrow band of 54–59% accuracy and 55–62% AUC-ROC. Contrary to the hypothesis that more expressive models would dominate, the ANN recorded the lowest performance overall (AUC = 55.75%), demonstrating that model complexity does not guarantee superior results on structured tabular data with limited feature counts. Feature importance analysis identifies betting market odds (moneyLine, projectedRuns) as the dominant predictors, followed by engineered rolling run features.

This study demonstrates that a unified, multi-source dataset combined with a consistent evaluation framework provides meaningful and reproducible insights into

UNIVERSITAS
GADJAH MADA
**Comparative Evaluation of Machine Learning and Deep Learning Models for Major League Baseball Game
Outcome Prediction**
Taichi Mori, Drs. Edi Winarko, M.Sc., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

predictive model performance for sports analytics tasks. The proposed framework contributes a modular and extensible methodology that can be adapted to other team sports such as soccer and basketball.

Keywords: Major League Baseball, sports analytics, machine learning, deep learning, game outcome prediction, comparative study