Simulating Blackjack: Exploring House Edge and Card Counting

Seán Houraghan (24209056) Deergh Kataria (24218299)
University College Dublin, School of Mathematics
ACM40960 — Projects in Maths Modelling

Study Aim

- Develop a blackjack simulator using Python
- Run Monte Carlo simulations using the simulator to explore how different playing strategies, game rules and card counting influences house edge
- Develop a model to predict whether or not a player is card counting based on action decisions, bet size and true count

Blackjack Basics

Blackjack is a popular casino card game where players aim to beat the dealer by getting cards totalling closer to 21 without going over (called "busting").

Card Totals:

- Number cards (2-10): Face value
- Face cards (Jack, Queen, King): Worth 10 points
- Ace: Worth 1 or 11 (whichever is better)

Gameplay: Each player receives two cards initially, as does the dealer (one face-up, one face-down). Players can "hit" (take another card) or "stand" (keep current total). Other options include "double down" (double bet, take one more card) or "split" (separate identical cards into two hands).

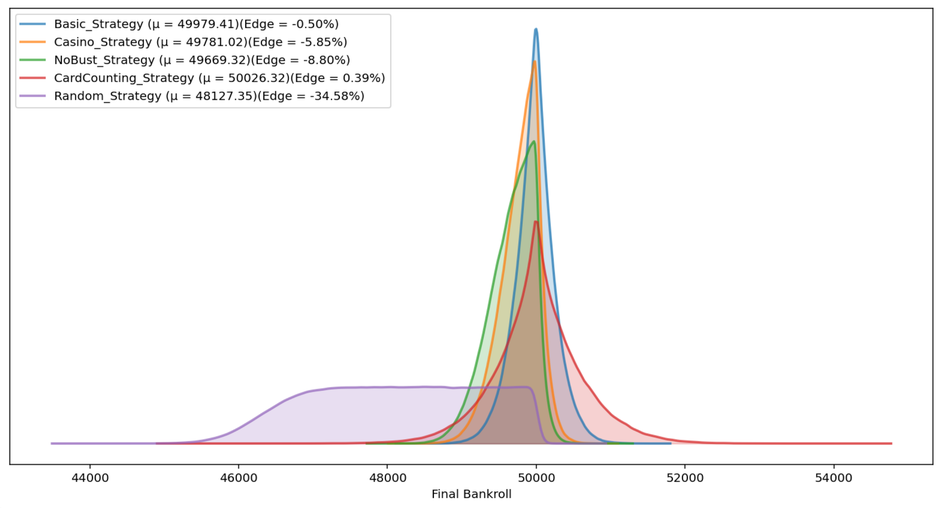
Winning:

- Get 21 with first two cards (Natural Blackjack)
- Get closer to 21 than dealer without busting
- Dealer busts whilst you don't

House Edge: House edge represents the statistical advantage a casino holds over players in gambling games, expressed as a percentage of each bet. It ensures long-term profitability for casinos regardless of short-term outcomes. For example, a 2% house edge means the casino expects to keep $2 from every $100 wagered over time.

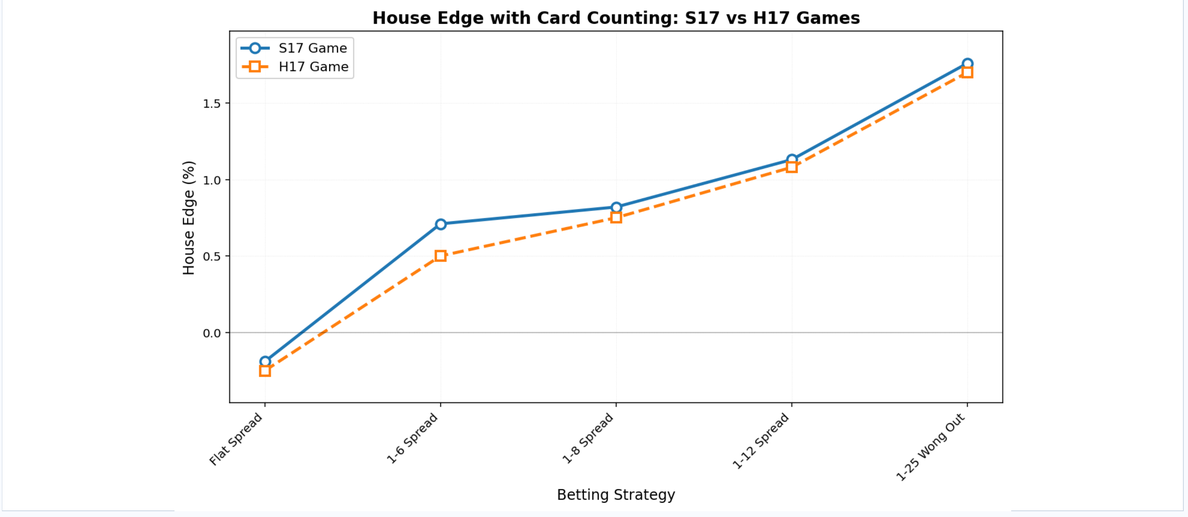
Basic Strategy: The mathematically optimal way to play every blackjack hand, developed through computer analysis of millions of hands. It tells players when to hit, stand, double, or split based on their cards and the dealer's upcard, minimising the house edge.

Playing Strategy Analysis



This density plot compares final bankroll distributions across five blackjack strategies. Card counting shows the only positive house edge (0.39%) with the highest mean bankroll, while random strategy performs worst (-34.58% edge). Basic, casino, and no-bust strategies cluster around similar means with negative house edges, demonstrating card counting's superior profitability.

Betting Strategy Analysis



This plot compares house edge percentages across different betting strategies for S17 and H17 blackjack games. Both game types show negative house edges (player advantage) with flat betting, transitioning to positive house edges with spread betting. S17 rules consistently provide slightly better conditions for card counters than H17 rules across all strategies.

Ruleset Impact

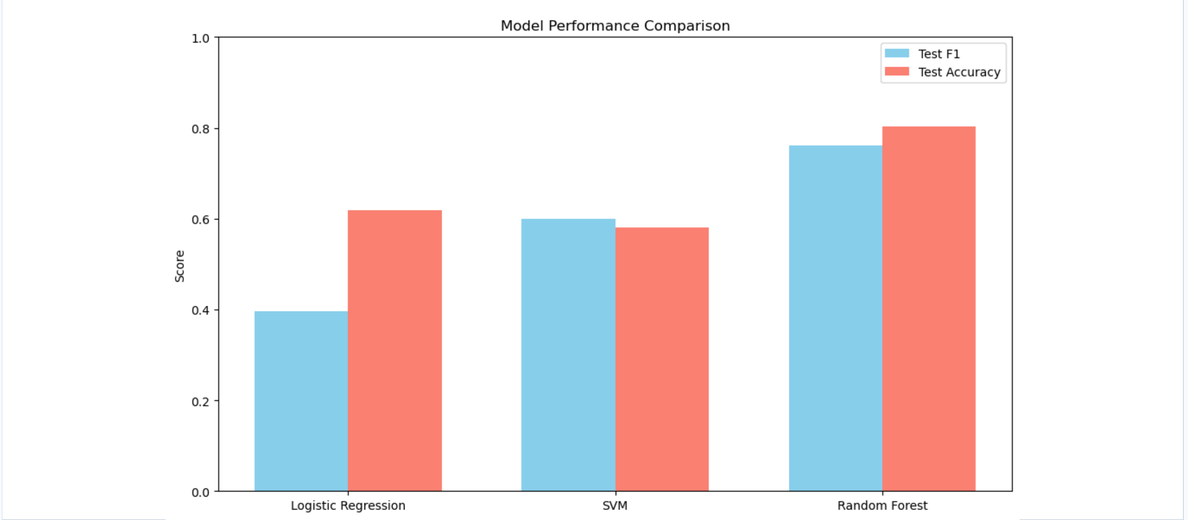
Basic Strategy	0.55%
Rule Change	Difference
1 Deck	+0.5%
2 Decks	+0.3%
4 Decks	+0.08%
6 Decks	+0.02%
Stand Soft 17	+0.18%
No Double After Split	-0.03%
No Late Surrender	-0.03%
6:5 Blackjack Payout	-1.18%

This table shows how rule variations affect the house edge in blackjack compared to basic strategy (0.55%). Multiple decks slightly increase the house edge, while soft 17 rules favour the house (+0.18%). Player-friendly rules like no double after split, no late surrender, and reduced blackjack payouts (6:5) significantly worsen player odds.

Variance

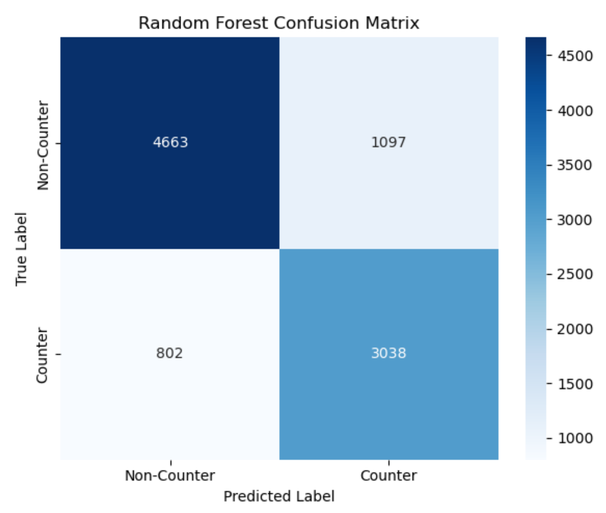
These results were achieved through the simulation of 5 million hands in total (10,000 trails of 500 hands). When playing basic strategy, the house edge has a 95% confidence interval of [-9.23, 8.13]. With such a wide interval, for card counters to be profitable they must play a lot to overcome variance.

Model Performance Comparison



The Random Forest model outperforms both SVM and Logistic Regression in terms of F1 score and accuracy, demonstrating superior capability for detecting card counters in the blackjack simulation. This makes it the optimal choice for the classification task.

Random Forest Confusion Matrix



The confusion matrix shows strong performance in distinguishing card counters (True Positives: 3038) from non-counters (True Negatives: 4663). While False Positives (802) and False negatives (1097) exist, the model achieves reliable detection overall, with room for refinement in edge cases.

Key Findings & Implications

Detection Feasibility:

- Card counters **can be detected** using gameplay data, though detection is not perfect.
- Random Forest performed best (**F1: 0.76**, **Accuracy: 0.80**); Logistic Regression and SVM were slightly less effective.

Most Predictive Features:

- **Bet Size** — strongest indicator; card counters vary bets based on true count.
- **True Count** — aligns with the theoretical basis of card counting.
- Secondary features: Player Value, Dealer Upcard, Action, Pair, Soft hand status.

Limitations:

- Simulated dataset — may not fully represent real casino play.
- Possible overfitting to simulated patterns.
- Class imbalance addressed via weighting but still a consideration.

Implications:

- Casinos can use ML models to flag potential card counters based on betting and hand patterns.
- Ensemble methods like Random Forest capture complex, non-linear behaviors more effectively than linear models.

Key Findings & Implications

Strategy Performance: Card counting is the only profitable strategy, achieving a positive house edge of 0.39% compared to basic strategy's -0.50%. Random strategy performs worst at -34.58%, demonstrating the critical importance of mathematical decision-making.

Betting Spreads: Progressive betting spreads significantly improve card counting profitability. The 1-25 Wong Out strategy yields the highest advantage (1.76% for S17 game), while flat betting results in negative returns despite card counting.

Variance Impact: Simulations of 5 million hands reveal the wide confidence intervals inherent in blackjack, emphasising that card counting requires substantial bankroll management and long-term play to overcome short-term variance.

GitHub Project Repository

Access the complete source code, data, and documentation for this blackjack simulation project on GitHub.