A Little Can Go A Long Way

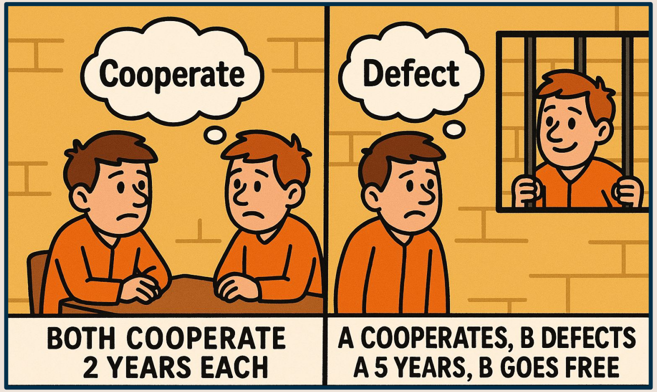
Classification, Prediction, and Ethical AI Strategy
Applications in Game Theory

Ritwik Awasthi

Prisoner’s Dilemma

Two isolated suspects must each choose to either **cooperate** or **defect**:

- If both cooperate, they each receive a light sentence
- If one defects while the other cooperates, the defector goes free and the cooperator receives a harsh sentence
- If both defect, they both receive a moderate sentence.



Iterated Prisoner’s Dilemma

Attempt to quantify or “score” someone's performance in a repeated prisoner's dilemma.

We assign and track scores, iterating continuously over several rounds. Player with the most score **“Wins”**.

Table 1. Payoff Matrix for the Prisoner’s Dilemma

Player A \ Player B	Cooperate (C)	Defect (D)
Cooperate (C)	(3, 3)	(0, 5)
Defect (D)	(5, 0)	(1, 1)

How does this apply to the real world?

Arms races: Two countries would both be safer (and save money) if they disarmed or limited weapons, but each fears the other will rearm and thus chooses to build more arms.

Business pricing: Competing firms would all earn higher profits if they kept prices high, but each has an incentive to undercut the other for market share, which often drives prices (and profits) down for everyone.

Environmental agreements: Nations benefit jointly if everyone reduces emissions, but any one country can gain a short-term economic edge by continuing to pollute, leading to higher overall emissions when each pursues that temptation.

Advertising and marketing: Companies could save on marketing costs if all agreed not to advertise heavily, but each firm fears losing visibility, so they all spend on ads, inflating costs without necessarily increasing collective sales.

Public goods and tax compliance: Individuals benefit if all pay their fair share of taxes (funding infrastructure, education, etc.), but each person has an incentive to evade taxes, relying on others to pay—if too many cheat, public services suffer.

Vaccination decisions: In communities where most people are vaccinated, disease risk is low; however, some may skip vaccination (to avoid cost or side effects), relying on herd immunity. If enough do this, herd immunity collapses and everyone is at higher risk.

Machine Learning Applications:

Classifying Player Strategies & Building an AI Player

What really is a “Strategy”

Grim-Trigger

Always Cooperate

Tit-For-Tat

Always Defect

Pavlovian

What do we already know?

Studies by Axelrod (1981) indicate that TFT is highly effective due to its simplicity, robustness, and reciprocity. Subsequent research highlighted vulnerabilities of pure TFT, leading to variants like Generous Tit-for-Tat and Suspicious Tit-for-Tat, incorporating elements of forgiveness and caution.

Prog.	TFT	T&C	NY	GR	SH	S&R	FR	DA	GR	DO	FE	JO	TU	NA	RAN	Mean	Rank Point	No. of Wins	Rank Wins
TFT	600	595	600	600	600	595	600	600	597	597	280	225	279	359	441	**504**	1	0	15
T&C	600	596	600	601	600	596	600	600	310	601	271	213	291	455	573	**500**	2	11	2
NY	600	595	600	600	600	595	600	600	433	158	354	374	347	368	464	**486**	3	1	13.5
GR	600	595	600	600	600	594	600	600	376	309	289	236	305	426	507	**482**	4	4	6
SH	600	595	600	600	600	595	600	600	348	271	274	272	265	448	543	**481**	5	3	11.5
S&R	600	596	600	602	600	596	600	600	319	200	252	249	280	480	592	**478**	6	10	3.5
FR	600	595	600	600	600	595	600	600	307	207	235	213	263	489	598	**473**	7	6	8
DA	600	595	600	600	600	595	600	600	307	194	238	247	253	450	598	**472**	8	4	9.5
GR	597	305	462	375	348	314	302	302	588	625	268	238	274	466	548	**401**	9	5	9.5
DO	597	591	398	289	261	215	202	239	555	202	436	540	243	487	604	**391**	10	6	6
FE	285	271	426	286	297	255	235	239	274	704	246	236	272	420	467	**328**	11	12	3.5
JO	230	214	409	237	286	254	213	252	244	634	236	224	273	390	469	**304**	12	10	1
TU	284	287	415	293	318	271	243	229	278	193	271	260	273	426	478	**301**	13	6	6
NA	362	231	397	273	230	149	133	173	187	133	317	366	345	413	526	**282**	14	2	11.5
RAN	442	142	407	313	219	141	108	137	189	102	360	416	419	300	450	**276**	15	1	13.5

doi:10.1371/journal.pone.0134128.t002

Is Tit-for-Tat the Answer? On the Conclusions Drawn from Axelrod's Tournaments - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Tournament-Scores-in-Axelrods-First-Tournament_fig2_281107220 [accessed 2 Jun 2025]

Dataset

60-20-20 Train Test Validation

Classification

Game	Round 1	Round 2	...	Round N	Strategy_1	Strategy_2
1	(C, D)	(D, C)	...	(C, D)	TFT	Random
2	(D, D)	(C, C)	...	(D, C)	TFT	Random
3	(C, C)	(C, D)	...	(D, D)	TFT	TFT

Prediction

Game	Round 1	Round 2	...	Round N	Payoff_1	Payoff_2
1	(C, D)	(D, C)	...	(C, C)	10	45
2	(D, D)	(C, C)	...	(D, C)	35	20
3	(C, C)	(C, D)	...	(D, D)	34	13

Classification Results

XGBoost

Table 4. Classification Report (5000 x 25 Rounds)

	precision	recall	f1-score	support
Tit-for-Tat	1.00	1.00	1.00	2000
Tit-for-2-Tat	0.88	0.96	0.92	2000
Suspicious-Tit-for-Tat	0.99	0.89	0.94	2000
Grim Trigger	0.99	1.00	1.00	2000
Pavlov	0.93	0.97	0.95	2000
Always Cooperate	0.99	0.97	0.98	2000
Generous TFT	1.00	1.00	1.00	2000
Soft Majority	1.00	0.97	0.99	2000
Random Strategy	0.90	0.81	0.85	2000
Alternator	0.73	0.75	0.74	2000
Gradual	0.74	0.73	0.73	2000
Limited Retaliation	0.98	0.98	0.98	2000
Tester	0.92	1.00	0.96	2000
accuracy			0.93	26000
macro avg	0.93	0.93	0.93	26000
weighted avg	0.93	0.93	0.93	26000

Model tuned with GridSearchCV

LSTM

Table 5. Classification Report

	precision	recall	f1-score	support
Tit-for-Tat	1.00	1.00	1.00	2000
Tit-for-2-Tat	0.92	0.91	0.92	2006
Suspicious-Tit-for-Tat	0.97	0.99	0.98	1942
Grim Trigger	1.00	1.00	1.00	2008
Pavlov	0.97	0.95	0.96	2052
Always Cooperate	0.99	0.99	0.99	1990
Generous TFT	1.00	0.99	0.99	2022
Soft Majority	0.98	1.00	0.99	1963
Random Strategy	0.86	0.87	0.87	1959
Alternator	0.71	0.78	0.74	1828
Gradual	0.80	0.73	0.76	2173
Limited Retaliation	0.99	0.99	0.99	2002
Tester	1.00	0.97	0.99	2055
accuracy			0.94	26000
macro avg	0.94	0.94	0.94	26000
weighted avg	0.94	0.94	0.94	26000

Model tuned with KerasTuner

Classification Performance

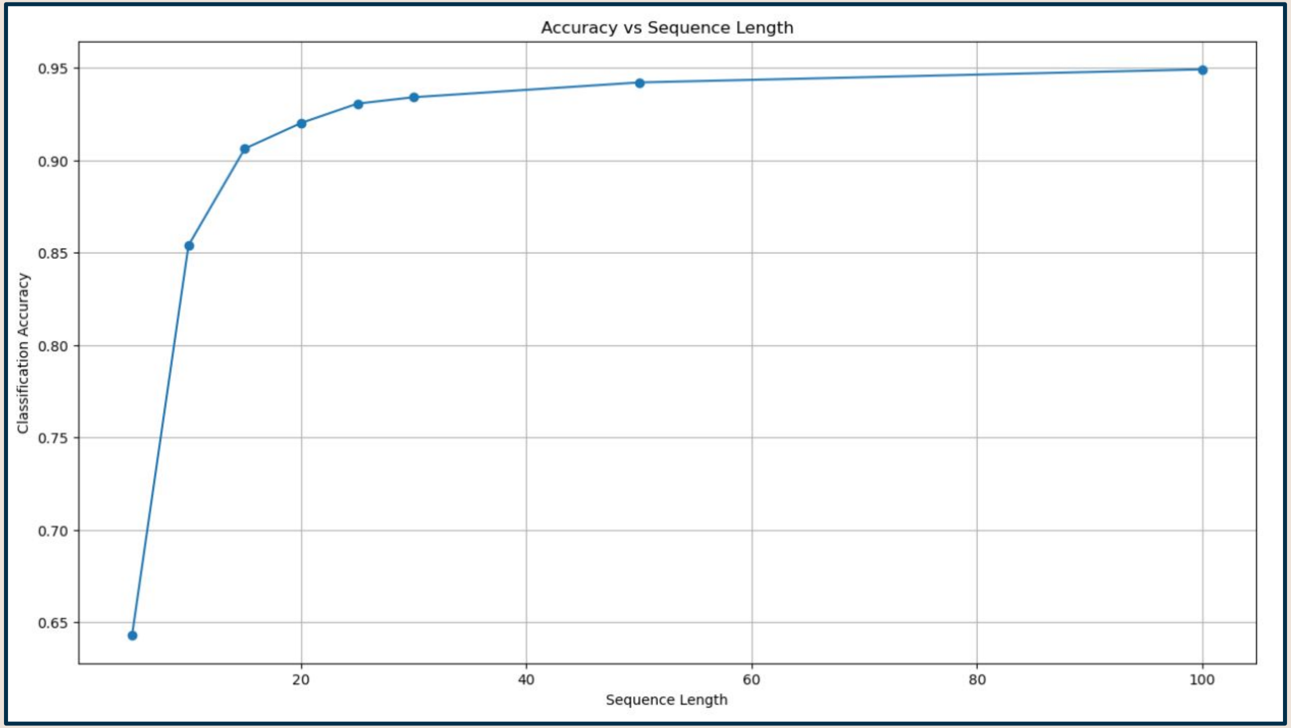
Overall accuracy of greater than 93%

Near-perfect precision/recall for TFT, Grim Trigger, Generous TFT

Lower scores for Alternator, Gradual due to complex behaviors

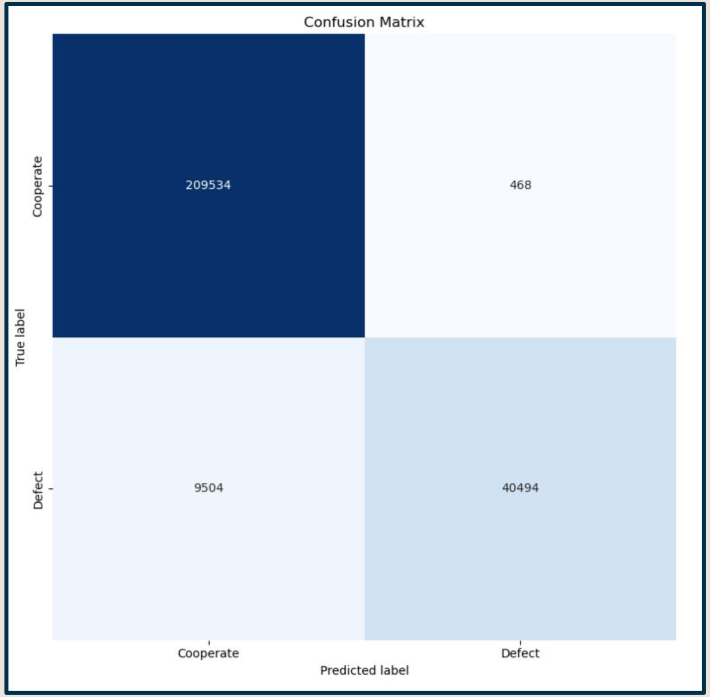
How many rounds do we need?

XGBoost - Accuracy as more rounds are played



Prediction Results

XGBoost



Model tuned with GridSearchCV

Table 6. Classification Report

	precision	recall	f1-score	support
0	0.9566	0.9978	0.9768	210002
1	0.9886	0.8099	0.8904	49998
accuracy			0.9616	260000
macro avg	0.9726	0.9038	0.9336	260000
weighted avg	0.9628	0.9616	0.9601	260000

0: **Cooperate** 1: **Defect**

Exploiting Predictability

Winning by extremes will lead to resentment and disdain.

Can we build an ethical player that ensures a "**win**" without completely destroying the enemy?

By mimicking the opponents move, we level the playing field: ensuring both players receive the same payout. This strategy banks on the fact that we can gain an "advantage" in the starting stages of the game, when our model is not prepared to make predictions, by implementing a Tit for Tat or similar strategy. Once we have enough data to predict an opponents move, we can level out the playing field by mimicking.

```
def policy(history, player):
    hist = history[-L:] if len(history) >= L else [('C','C')] * (L - len(history)) + history
    feat = []
    for p1, p2 in hist:
        feat.extend([1 if p1 == 'D' else 0, 1 if p2 == 'D' else 0])
    prob_comply = clf.predict_proba([feat])[0][1]
    return 'C' if prob_comply > 0.5 else 'D'
```

Tournament With An AI Player

10 games per distinct pair of strategies.

Games will have varying number of rounds to exhaust model performance.

For each tournament, we will record number of "Wins" and "Average Payoff" to gauge model performance.

Table 7. Average Payoffs (Rounds per Match: 50)

Strategy	Average Payoff
Tit-for-Tat	1445.54
Generous TFT	1441.08
Tit-for-Two-Tats	1423.31
Soft Majority	1408.00
Grim Trigger	1403.85
Gradual	1386.38
Pavlov	1374.62
Limited Retaliation	1369.38
Alternator	1296.31
Always Cooperate	1267.85
Tester	1237.85
Random	1223.15
Suspicious TFT	1167.00
AI	1089.54

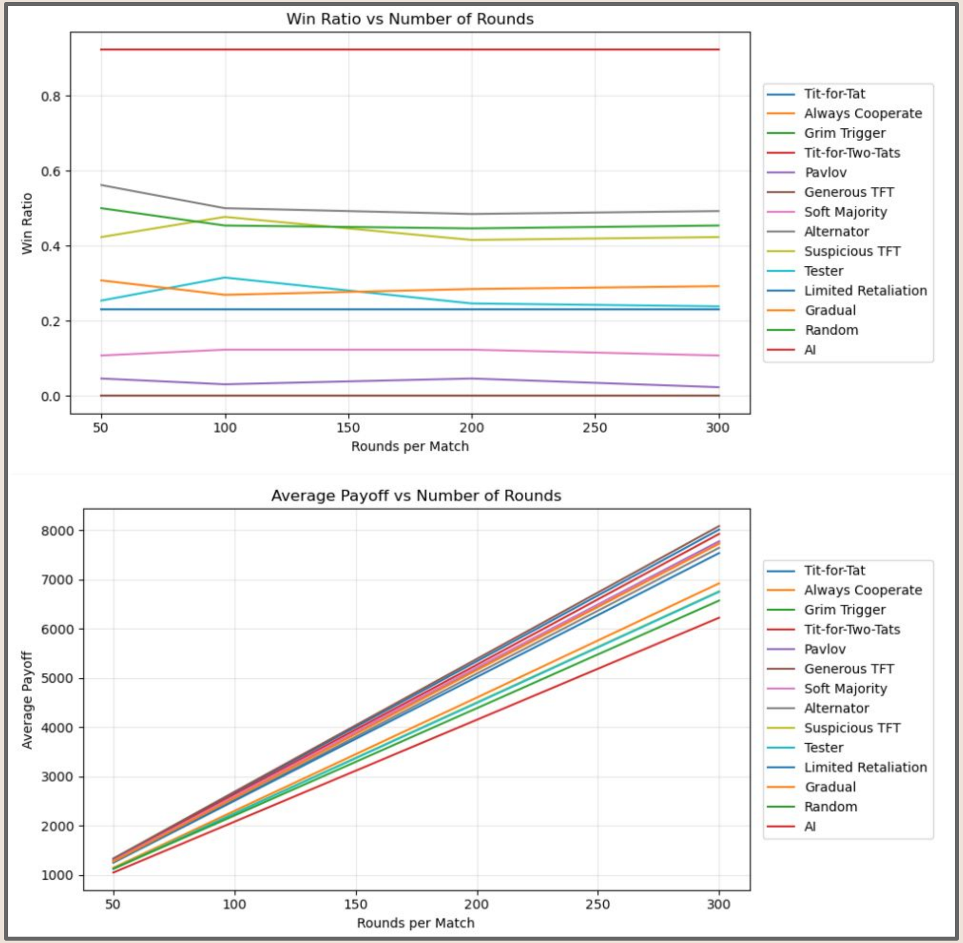
Table 8. Win Counts (Out of 130 Games)

Strategy	Wins / Games
AI	120 / 130
Alternator	77 / 130
Random	66 / 130
Suspicious TFT	54 / 130
Gradual	37 / 130
Tester	32 / 130
Grim Trigger	30 / 130
Limited Retaliation	30 / 130
Soft Majority	15 / 130
Pavlov	2 / 130
Tit-for-Tat	0 / 130
Always Cooperate	0 / 130
Tit-for-Two-Tats	0 / 130
Generous TFT	0 / 130

Results

AI player achieves the highest Win Rate with lowest average payoff across every number of rounds tested per game.

This showcases the possibility of applying Machine Learning to create ethical players with a system that can be model by the iterated prisoner's dilemma.



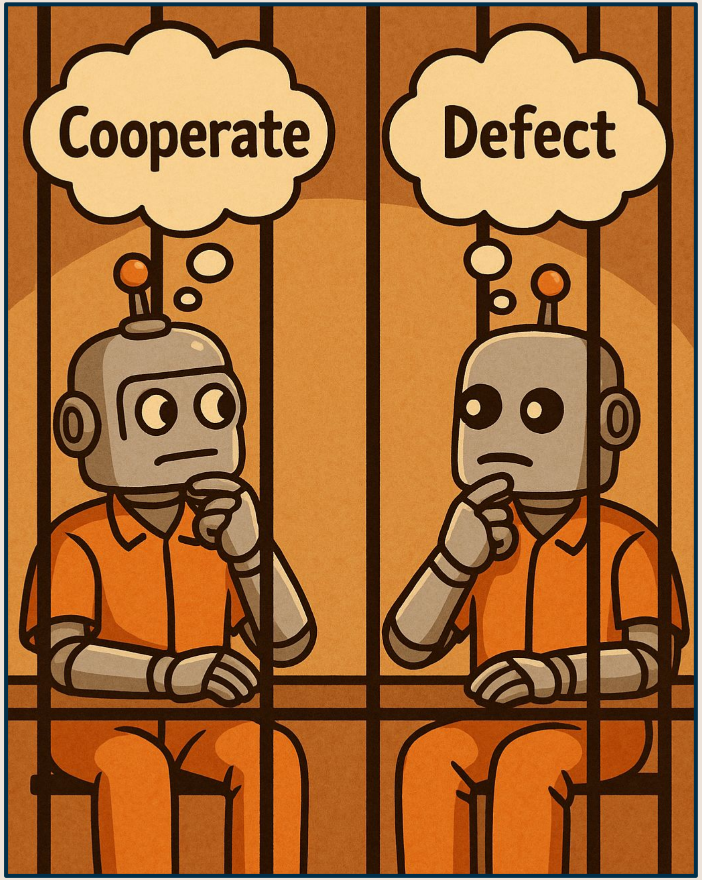
Conclusion

Conclusions

- Classifier identifies strategy ideally after 25 rounds.
- Predictive model achieves near perfect forecasting on opponent defections.
- AI wins majority of matches while minimizing payoffs across varying round lengths per game.

Future Work

- Incorporate noise to test robustness.
- Simulate multi-agent ecosystems.
- Analyze the effect of using different starting strategies.
- Analyze how AI players play against each other



References

- Axelrod, R. (1984). The Evolution of Cooperation. Basic Books.
- Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. Science, 211(4489), 1390–1396.
- Nowak, M. A., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. Nature, 364(6432), 56–58.