Cheat Detection in Chess: Analyzing PGN Files

Dissertation Presentation

BITS ID – 2022AC05327

Ankita Pal, WILP BITS Pilani – MTech in AI & ML February 2025

Supervisor: Milin Shah, VP and Technology Manager at Bank of America

Presentation overview

1. Introduction to Chess Cheating	7. Results & Statistical Analysis
2. Problem Statement & Significance	8. Challenges & Limitations
3. Research Objectives & Contributions	9. Conclusions & Key Findings
4. Literature Review & Current Approaches	10. Future Work & Enhancements
5. Methodology Framework	11. Discussion
6. Implementation Details	12. Q&A

01

INTRO

INTRODUCTION TO CHESS CHEATING

Cheating in chess on online platforms is on a constant rise. There’s been an increased prevalence since 2020. Engine assistance: Stockfish, Leela Chess Zero & AlphaZero provide superhuman move accuracy

Detection challenges

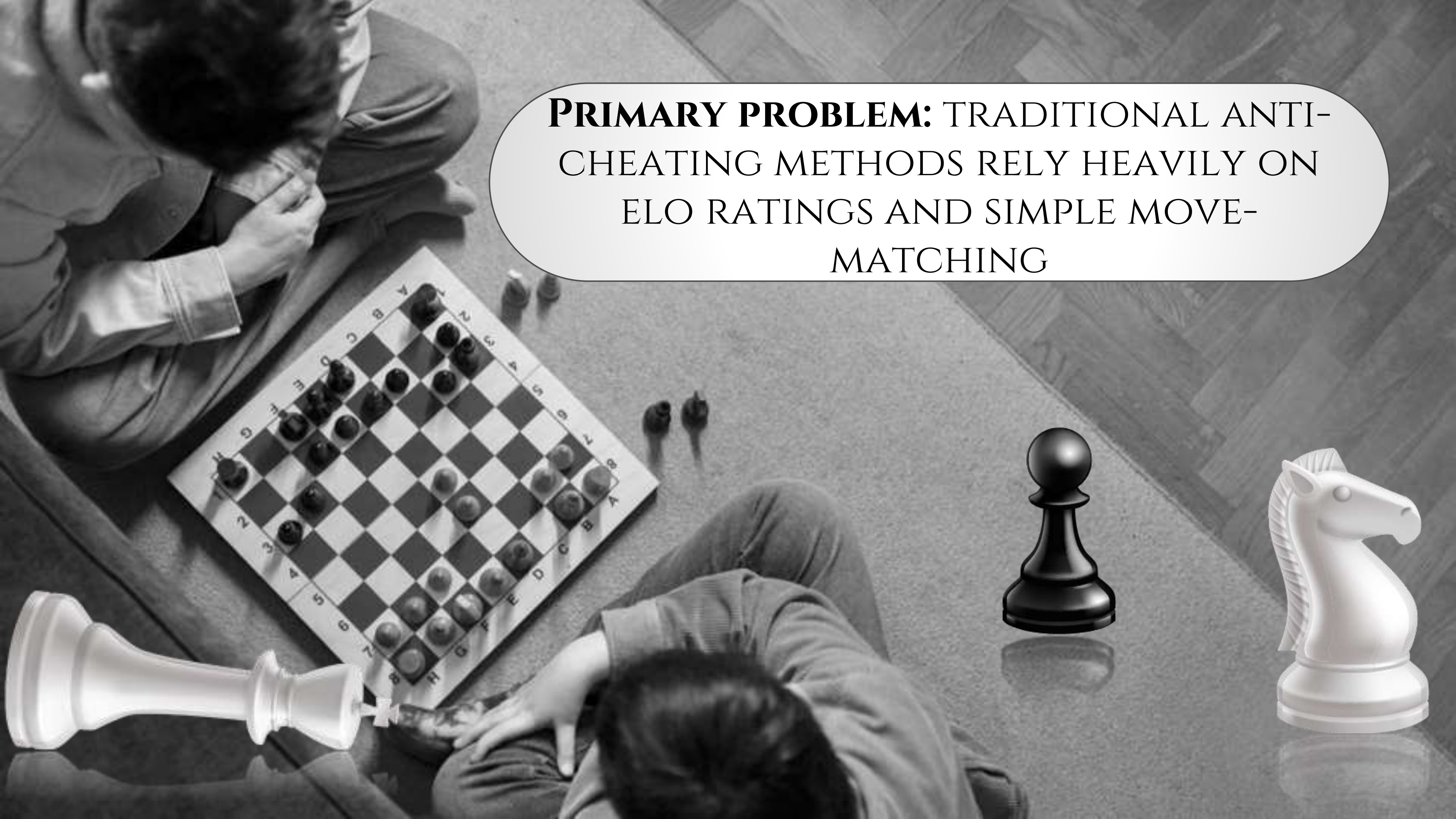
- No physical tells in online environments
- Sophisticated methods (toggling between screens, consulting engines selectively)
- Varying skill levels making baseline difficult to establish

Impact on Chess Community: Erosion of trust and competitive integrity

Need for automated detection: Statistical methods to complement human arbitration

02

Problem statement


Primary problem: traditional anti-cheating methods rely heavily on elo ratings and simple move-matching

Limitations of current approaches:

- Proprietary “black box” algos
- High false positive rates
- Difficulty in distinguishing coincidental engine alignment from cheating

Research question: Can PGN analysis identify statistical anomalies indicative of computer assistance?

Significance: Maintaining competitive integrity in online patforms, supportive tournament arbiters with objective evidence, providing transparent detection methods

03

Research objectives & contributions

Primary objectives

01 — Develop automated framework for cheat detection using PGN data

02 — Implement CPL metrics to evaluate move precision

03 — Create visualization tools for pattern identification

04 — Build machine learning models to classify suspicious play

Novel contributions:

- Integration of multiple detection parameters (not just engine matching)
- Focus on game-length analysis as a complementary indicator
- Open-source approach with transparency in methodology
- Use of modern ML techniques (achieving 95% accuracy with Neural Networks)



What is CPL (Centipawn Loss)?

Definition: Numerical measure of move quality relative to optimal engine move

·100 centipawns = value of one pawn
·Lower CPL = closer to perfect (engine) play

CPL Characteristics:
·Grandmasters: ~15–20 average CPL
·Club players: ~50–100 average CPL
·Beginners: ~150+ average CPL

Detection Relevance:
·Unnaturally consistent low CPL across games suggests engine assistance
·Important to consider game complexity, time controls, and position criticality
·Pattern of CPL more revealing than absolute values

04

Literature review

Literature review

- **Historical Development**:
- Early statistical approaches (Regan & Haworth, 2011)
- Chess.com's proprietary system development (2014–present)
- Lichess's IRWIN detector (2017)

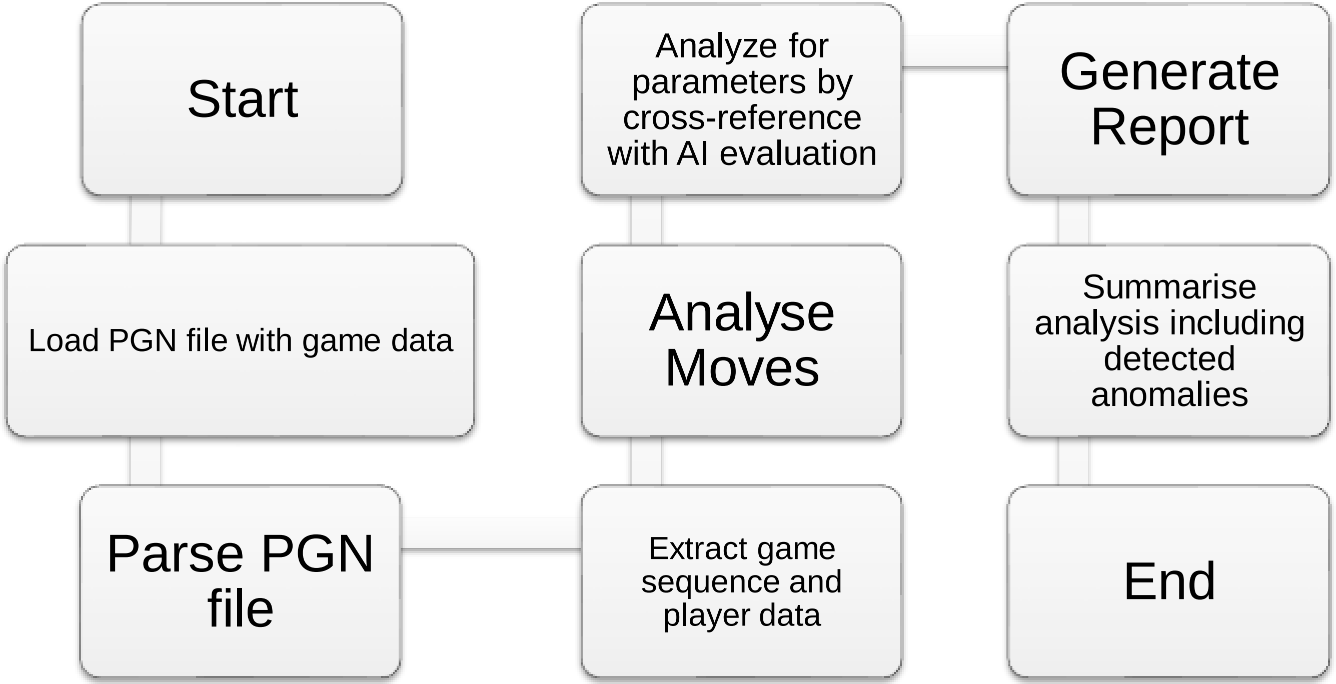
- **Current Detection Methods**:
- Z-score analysis of move matching percentages
- Time usage patterns and correlation with move difficulty
- Move consistency across different game phases
- Player performance deviation from historical results

- **Research Gaps**:
- Limited transparency in commercial implementations
- Inadequate contextual move analysis
- Challenges in real-time monitoring and intervention
- Scarcity of labeled datasets for supervised learning

05

Methodology framework

PGN Analysis Methodology



Process flow

01	Game data	Extract game data and player information
02	COMPUTE	Compute move-by-move engine metrics
03	CPL	Apply CPL and other metrics
04	Analysis	Generate comprehensive analysis report

06

IMPLEMENTATION

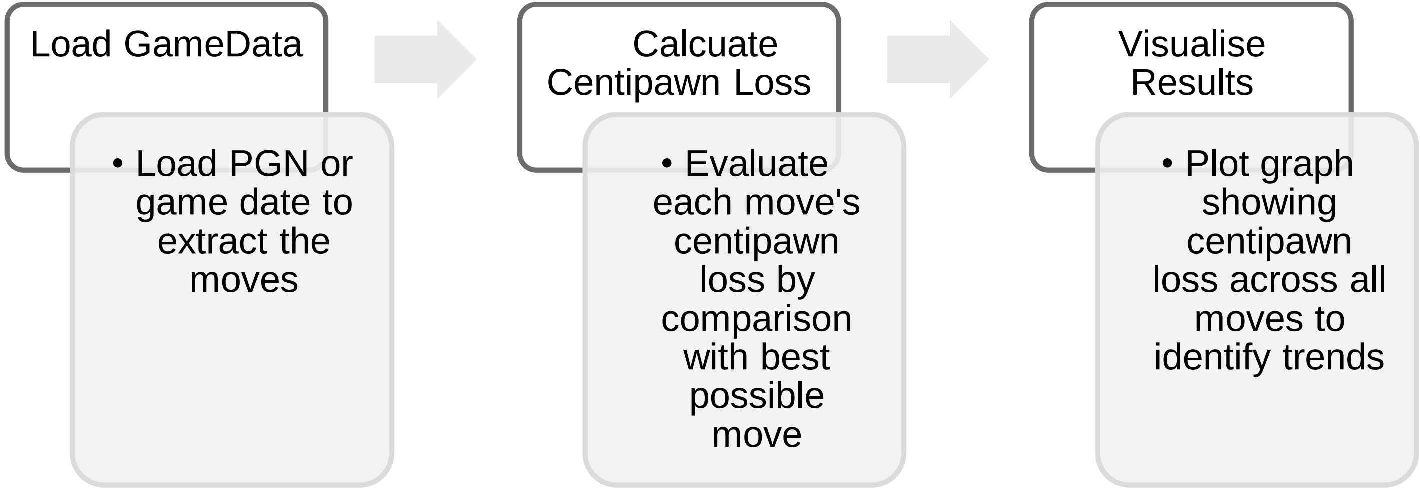
Implementation

Dataset Composition:
- PGN files from Chess.com and Lichess
- Range of Elo ratings (800-2500)
- Mix of known fair games and confirmed cheating cases
- Various time controls (bullet, blitz, rapid, classical)

Technology Stack:
- Python with chess.pgn library for parsing
- Stockfish 16 for engine analysis (depth 20)
- Pandas for data manipulation
- Matplotlib and Seaborn for visualization
- Scikit-learn and TensorFlow for ML models

Implementation Challenges:
- Handling large PGN files efficiently
- Optimizing engine analysis time
- Accounting for different time controls in evaluation

Centipawn Loss Across Moves



Centipawn loss calculation process

Implementation Details:

•Extraction of move sequences in UCI format

•Engine evaluation at consistent depth (depth 20)

•Calculation formula: CPL = max(0, evaluation_before_move – evaluation_after_opponent_best_response)

•Exclusion of book moves (first 10 moves) to reduce noise

•Special handling for forced moves and time trouble positions

Machine learning model performance

Algorithm	Accuracy	Precision	Recall	F1-Score
Decision Tree	85%	88%	83%	85%
Random Forest	92%	90%	93%	91%
Neural Network	95%	94%	96%	95%

Feature Engineering:
- Average CPL across game phases
- CPL variance and standard deviation
- Move matching percentage with top engine choices
- Time usage patterns and correlation with move complexity
- Game length relative to player rating

Model Selection Rationale:
- Neural Network provides best overall performance
- Random Forest offers good interpretability with strong results
- Decision Tree serves as baseline and provides explainable rules

07

Results & analysis

Results: identifying suspicious patterns

Primary indicators

- Consistently low CPL across complex positions
- Unnatural CPL stability
- High alignment with top engine choices in critical positions
- Incongruence between move strength and player rating

Secondary indicators

- Unusual game length patterns
- Inconsistent time usage (very quick complex moves)
- Performance spikes in specific tournaments or time periods
- Drastic improvement without corresponding rating increase

Classification thresholds

Low confidence

Only secondary indicators present

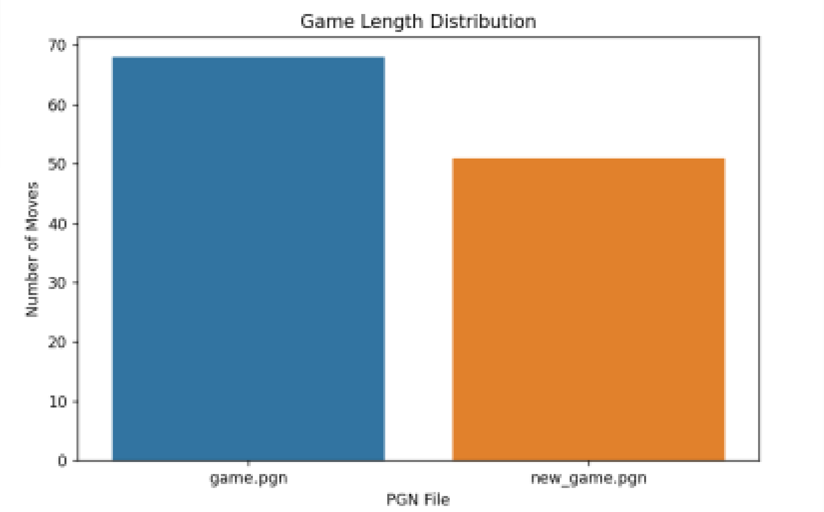
Medium confidence

One primary + multiple secondary indicators

High confidence

Multiple primary indicators present

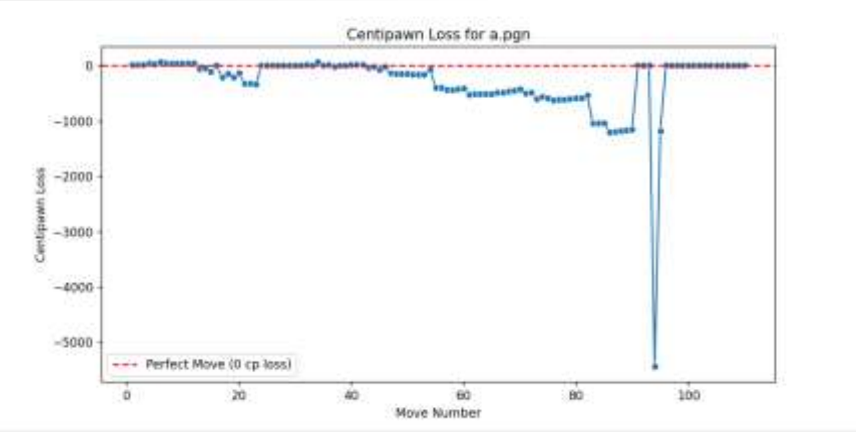
Game length distribution analysis

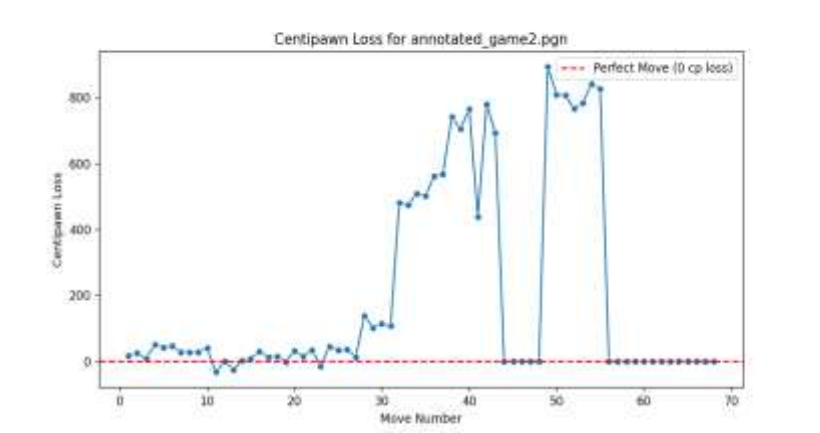


Analytical Insights:

- Normal distribution centered around 40 moves for fair play
- Bimodal distribution for suspicious accounts:
  - ➢ Very short games (quick wins with engine assistance)
  - ➢ Very long games (prolonged optimal defense)
- Correlation between game length and outcome for suspected cheaters
- Time control influence on distribution patterns

CPL Trend Analysis





Pattern Interpretation:

·Figure 1: Typical human player showing variable CPL with occasional spikes

·Figure 2: Suspected engine assistance showing unnaturally flat, low CPL

Key Observations:

·Critical position identification (where decisions are complex)

·Move timing correlation with CPL (suspiciously quick perfect moves)

·Consistency across multiple games from same player

·Phase-specific analysis (opening, middlegame, endgame)

08

Challenges & limitations

TECHNICAL CHALLENGES (TC) & METHODOLOGICAL LIMITATIONS (ML)

01

TC: Computational intensity of engine analysis

02

TC: Handling different chess engine versions/depths

03

TC: Accommodating various time controls in analysis

04

ML: Reliance on annotated PGN files for training data

05

ML: False positives with very strong human players

06

ML: Intentional suboptimal play to avoid detection

Ethical considerations

False accusation

Balancing detection with false accusation risk

transparency

Transparency in methodology to ensure fairness

privacy

Privacy concerns with player data analysis

09

Conclusions & key findings

Primary Conclusions

CPL analysis effectively differentiates fair from engine-assisted play

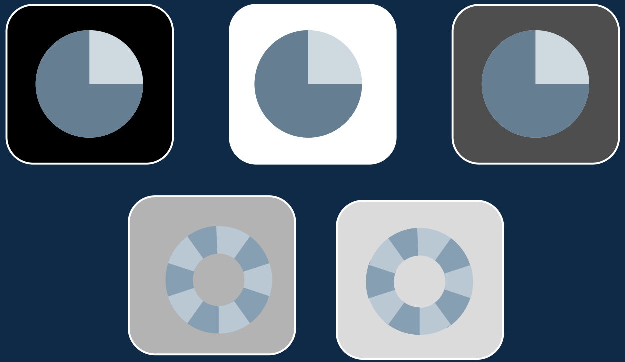
Game context (complexity, phase, time control) crucial for accurate detection

Multi-factorial approach reduces false positives

Neural Network models achieve 95% accuracy in classification

Key Findings

- Pattern consistency more revealing than individual move accuracy
- Game length distribution provides valuable supplementary evidence
- Time usage patterns strongly correlate with suspicious play
- Combined metrics approach outperforms single-factor detection



Practical applications

Tournaments

Tournament screening tool for arbiters

55%

Chess sites

Platform-level implementation for online chess sites

45%

Educational resource for fair play enforcement

10

Future work

Technical enhancements

- Real-time analysis capability for tournament monitoring
- Cloud-based implementation for scalability
- Mobile application for arbiters and tournament directors
- Integration with chess platform APIs

Improvements and Research Extensions

Methodological Improvements:
 - Expanded training dataset with more confirmed cases
 - Advanced neural network architectures (LSTM, Transformer)
 - Incorporation of psychological factors and playing style
 - Adaptive thresholds based on time control and player rating

Research Extensions:
 - Application to other strategic games (Go, Shogi)
 - Self-improving detection through continuous learning
 - Combined hardware/software solutions for tournament play

Acknowledgements

Supervision and Guidance

- Milin Shah, Bank of America (Primary Supervisor)
- BITS Pilani Department of AI & ML Faculty

Technical Support

- Stockfish Development Team
- Python-chess Library Contributors

Data Sources

- Chess.com Research Dataset
- Lichess Open Database

Research Assistance

- Chess Community Forum Contributors
- Fellow BITS Pilani Students & Researchers

Chess community friends

- Manan Pahwa, Purdue University (EECS), Indiana, USA
    i. PGN dataset contributions (chess.com) for test cases and sample games
    ii. Providing expert perspectives on human play patterns
    iii. Invaluable feedback on early detection algorithms
    iv. Dissertation idea incubator
    v. Chess coach

- Moin Memon, Bank of America, GBS Chess Championship Winner 2022
    i. Assistance with testing across various skill levels
    ii. Local chess club member participant for controlled testing
    iii. Tournament coordinator sharing insights on practical implementation needs

Acknowledgements continued

11 & 12

Discussion + Q&A

Panel members

Any feedback and questions for me?

THANK YOU

Any questions? Please reach out to me –

2022AC05327@wilp.bits-pilani.ac.in
+91 7490054195

BITS Pilani
Pilani | Dubai | Goa | Hyderabad
Work Integrated Learning Programmes