Computer Science
UNIVERSITY OF TORONTO

Microsoft

Bridging the Gap Between Human and Artificial Intelligence in Chess

Reid McIlroy-Young, Siddhartha Sen, Jon Kleinberg, Ashton Anderson

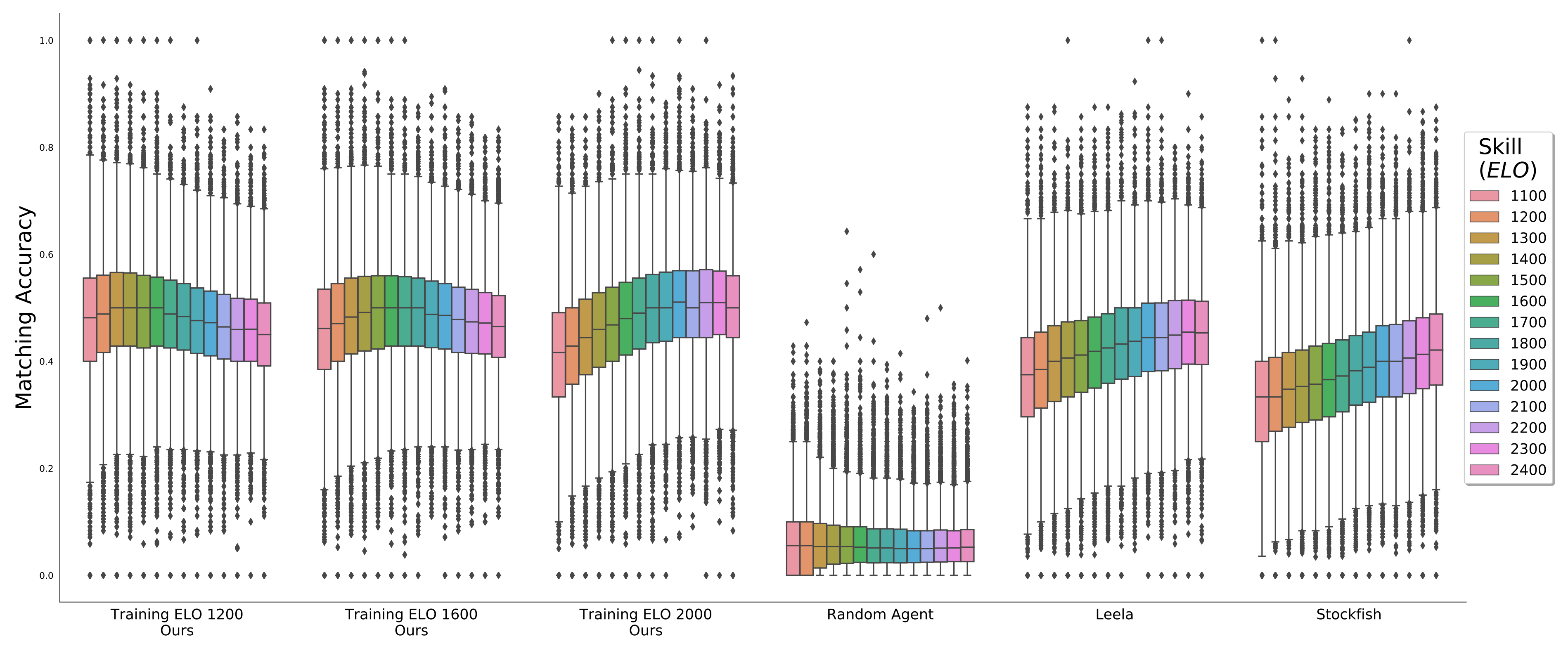
RQ1: Can we predict the next move a human will make?

RQ2: Can we predict when a human will make a mistake?

RQ3: Can we build an RL system that is risk adverse like a human?

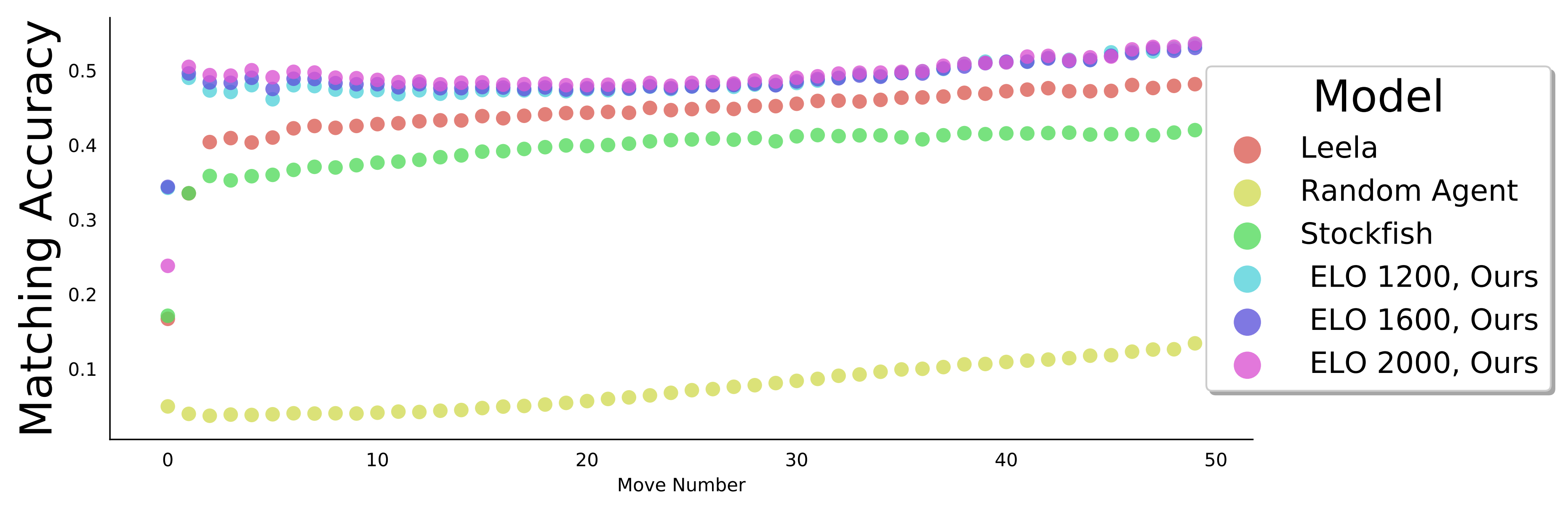
- Artificial systems can easily beat humans at chess, but cannot imitate them
- We believe that moving from superhuman to human will make better learning tools for humans and improve our artificial systems
- The machine learning components, (Monte Carlos Tree Search (MCTS) , ResNet, etc) used for chess generalize to other domains
- Non-ML superhuman chess engines (*Stockfish*) are also available for insight

Move Matching by Human Skill



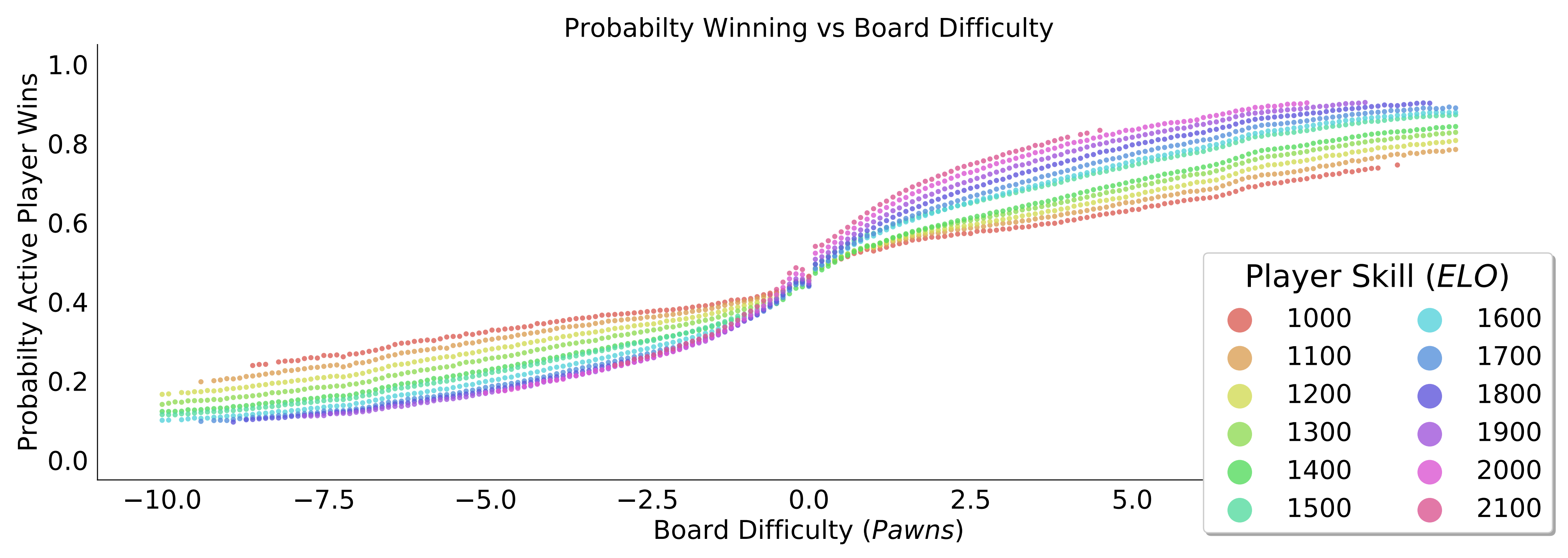
We selected 30,000 games from our holdout set from each ELO range and measured the percentage of moves the models matched the black player's. No MCTS roll outs were performed while *Stockfish* was allowed to search.

Move Matching by Move Number



We can also look at the percentage of moves the models matched at the n'th move of the game. Random agent shows an upward trajectory due to the number of legal moves decreasing.

Human Win Probability



100,000+ observations per point. We can thus give any board a win probability, and for any move the change in win probability. Also note better players are more advantaged by good positions.

Terms and Data

- **Board** A chess board with pieces, a single game state
- **Blunder/Mistake** we define a mistake as a move that lowers your win probability by more than 10% (about 9% of moves)
- **ELO** a measure of player skill, average is 1500, higher is stronger
- 850 million human games from *lichess.org*, going back to 2013
- Games have information about player skill, format and time
- We excluded moves or games with low time from most analysis

Model

- Based on *Leela Chess* an open source *AlphaZero* implementation
- Game boards are input as 8x8 images with each channel a bit mask of each type of piece (12 total), plus last 7 boards the same way
- 6 residual blocks each with 64 channels, then two fully connected layers
- Outputs are predicted next move (π) and expected reward (Q)

Training

- Trained only on human games with both player's ELOs within a range
- About 10 million games per ELO range, holdout set of the most recent month was used for testing
- Batch size of 2048 and trained in 140,000 steps
- Used momentum optimizer and learning rate of .02
- Training took 9 days on average on a *Tesla K80*

Results

- Our models outperform all others at matching the next move a human will make
- Engine performance weakly correlates with human move matching, our engines are the weakest in playing chess
- We can tune the models using information about the human's skill
- Deeper networks and allowing MCTS improve performance

Mistake Prediction with Decision Trees

- Given a chess board, what is the probability of the active player making a mistake?
- We trained a series of classifiers using simple features of the game state, ended with decision trees
- We care about the performance on the unbalance real world data (9% blunder rate)

	Accuracy	
Input Data	balanced	real
Just Time	0.58	0.35
" + Difficulty	0.65	0.50
Time + ELO	0.58	0.37
All the Above	0.64	0.53
" + Piece Counts	0.65	0.57

Mistake Prediction with Deep Learning

- Looking just at the board, can we predict if a human will make a mistake?
- Using 6 residual blocks with 64 channels looking at just a single board we get 66% accuracy
- Using 3 fully connected layers with 256 nodes we get 61% accuracy
- Not doing much better than the decision trees, but the input data are less
- Work is ongoing to improve accuracy, we hope to get above 80%

You can play against our model: *lichess.org/@/haibrid_bot* and code is at *github.com/CSSLab/haibrid-chess*

Safe Reinforcement Learning Search

- In chess superhuman engines are known for choosing risky lines even when safer ones exist
- If this behaviour can be controlled we could convert superhuman performance to human
- We have setup a set of 300+ board positions with an unsafe line (only one good next move) and a safe line (many good next moves), these form a testing set
- We have tried some modifications to the MCTS, but they either have no effect (picking the highest Q value) or cause a collapse (adding smoothing the policy distribution)
- Work is ongoing to add and test new modifications to MCTS