

Simulating Human-Level Performance in Competitive Pokémon Battling

1st Kaleb Conner
Department of Computer Science
Middle Tennessee State University
Murfreesboro, TN USA
kac9k@mtmail.mtsu.edu

2nd Collin Dyer
Department of Computer Science
Middle Tennessee State University
Murfreesboro, TN USA
cd4c@mtmail.mtsu.edu

3rd DaVonté Lewis
Department of Computer Science
Middle Tennessee State University
Murfreesboro, TN USA
drl3r@mtmail.mtsu.edu

4th Mathew Wells
Department of Computer Science
Middle Tennessee State University
Murfreesboro, TN USA
mjw6k@mtmail.mtsu.edu

Abstract—In recent years, neural networks have shown that machines are capable of learning and performing various tasks with a level of skill we never thought possible before. A number of trainee networks have even displayed the ability to play games of varying difficulties at human or superhuman proficiency. Pokémon is deceptively complex game with clear, discrete choices to be made, making it a great candidate for the implementation of machine learning. By having our network watch a thousand Pokémon battles, over and over, we were able to produce an agent capable of predicting the actions a human would take 90% of the time.

Index Terms—Pokémon, Neural Network, LSTM, Machine Learning,

I. INTRODUCTION

Pokémon is an adversarial, turn-based strategy battle simulator with immense global recognition and cultural staying power. The classic configuration consists of two players engaged in combat through the use of up to six “pocket monsters” (Pokémon) each. The inclusion of elements such as swapping, held items, hidden abilities, status effects, and Pokémon typing make for an incredibly dynamic and necessarily complex combat system. In this project, our goal was to train a neural network capable of achieving a win rate above 50% on Pokémon Showdown, a popular platform for competitive battling. Due to the sheer breadth of the Pokémon environment, we have limited the software’s scope to the top percentage of Pokémon used in competitive play, as well as their most common moves and held items based on Showdown data. Considering these are the teams and configurations encountered most frequently on Pokémon Showdown, our target platform, refining our team to fit within this structure was our primary objective. Towards that end, we employ teacher-forcing in combination with a Long Short-Term Memory (LSTM) recurrent neural network.

II. BACKGROUND

Strategic and competitive games have long been a training ground and general area of interests for the implementation of



Fig. 1. Pokémon Showdown is a web-based, free and open source Pokémon battle simulator

AI and neural networks. Their ability to explore numerous potential action paths in a relatively short span of time lends itself to strategic optimization, allowing them to surpass even human performance under the proper training conditions. Superhuman performance has been achieved using neural networks in various domains, including classic board games such as chess, shogi, and Go [1], as well as several Atari games [2]. In all instances, the aforementioned neural networks were the result of some variation of unsupervised, reinforcement-based learning, where a reward or punishment was either bestowed or withheld given some pre-established state configuration [3]. While effective, the amount of data and exploration necessary to produce a successful network implementing reinforcement learning is quite costly. Given the complexity of the Pokémon state space and our relatively limited access to relevant data, we instead opt for a significantly less computationally expensive supervised approach, where we test to see if a network can generalize human understanding and intuition in the Pokémon Showdown battle environment through mimicry alone.

III. METHODS

A. Data Collection

Our first task was the collection and configuration of viable input on which to train our neural network. Finding viable

training and testing data was relatively simple: Pokémon Showdown allows users to save battle replays; assuming they are able to access the active battle, this feature is available to any and all users on the site. The replays are kept on a separate but related website which provides details on how to scrape more viable data stored in a JSON format. The JSON component, however, is simply user metadata, with the actual replay being far more difficult to parse. In order to keep server request spam to a minimum, we scraped the saved replays and attempted to extract relevant information over the course of a two week period. This resulted in a cache of well over 3000 replays to convert into a structure suitable for network input. The replays were stored as easy-to-read text, consisting of numerous individual environmental factors to be parsed and tracked. The greatest challenge we faced concerned “volatile” status effects, a large class of effects applied to a player’s Pokémon that can disappear over time or when switched. The replay storage handles adding and removing automatically, there is, however, an exceedingly vast number of them to consider. By the end of parsing, there were over 50 flags that could be set. Ultimately, after successfully parsing the majority of replays, we decided to incorporate 207 parameters into our model, consisting of each Pokémon’s stats, known moves at a given turn, various field effects, and any stat buffs or debuffs.

Subsnooze Details Level: 50 Gender: Male HP: 255 Rb: No Items Choice Scarf Ability Haze Stats HP: 255 Atk: 140 Def: 140 SpA: 140 SpD: 140 Spc: 255	Subsnooze Details Level: 50 Gender: Male HP: 255 Rb: No Items Choice Scarf Ability Haze Stats HP: 255 Atk: 140 Def: 140 SpA: 140 SpD: 140 Spc: 255
Subsnooze Details Level: 50 Gender: Male HP: 255 Rb: No Items Choice Scarf Ability Haze Stats HP: 255 Atk: 140 Def: 140 SpA: 140 SpD: 140 Spc: 255	Subsnooze Details Level: 50 Gender: Male HP: 255 Rb: No Items Choice Scarf Ability Haze Stats HP: 255 Atk: 140 Def: 140 SpA: 140 SpD: 140 Spc: 255
Subsnooze Details Level: 50 Gender: Male HP: 255 Rb: No Items Choice Scarf Ability Haze Stats HP: 255 Atk: 140 Def: 140 SpA: 140 SpD: 140 Spc: 255	Subsnooze Details Level: 50 Gender: Male HP: 255 Rb: No Items Choice Scarf Ability Haze Stats HP: 255 Atk: 140 Def: 140 SpA: 140 SpD: 140 Spc: 255

Fig. 2. A significant amount of data associated with each individual Pokémon is required to produce adequate input for training

B. Model

For training, we chose an initial network built with a simple LSTM layer. The idea being that the network would be able to watch an entire replay while remembering every previous turn taken in a given battle. This approximates the amount of information and level of detail possessed by an average human player.

IV. RESULTS

Due to the limited amount of resources available to our primary training machine, our initial training and testing data consisted of only 1000 replays. Despite this, we were able to complete 3500 epochs over the course of an hour-long training session. Our first run was somewhat disappointing, holding below 20% accuracy until about the half-hour mark of training, when it jumped to a high of roughly 90%; this was sustained for the remaining epochs (Fig.3). The results of our second run were more promising, with our network holding above 90% throughout the entirety of training. We

saved both models using keras built in function and ran the second test through the rest of our data. This also gave a result above 90% with over 2000 previously unseen replays, indicating potentially promising performance in future real-world testing.

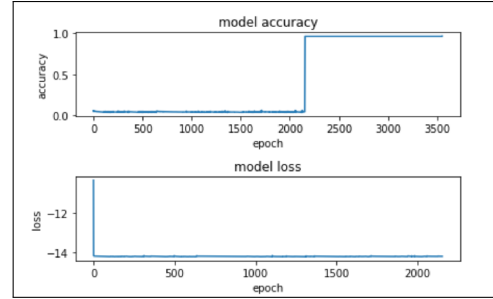


Fig. 3. Initial Network Performance

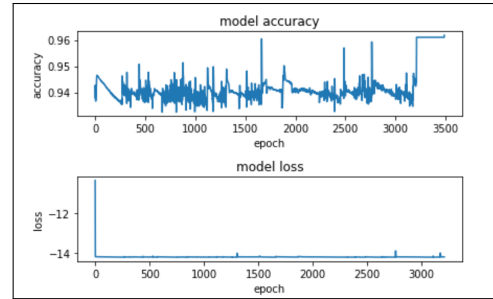


Fig. 4. Modified Network Performance

V. DISCUSSION

Further work is needed in this area. As of the writing of this paper, we have yet to attempt any real-world testing, however, we expect testing to go reasonably well if the results of our second round of training are any indication (Fig.4). Next steps would include the generation of fake replay data based on a live battle using one or both of our existing models. Without further testing and the full results of a live performance, we cannot know whether our networks, as they are currently constituted, would be capable of achieving a 50% win rate, approximating human-level play in competitive Pokémon battling. We look forward to building upon the strong foundation outlined thus far.

REFERENCES

- [1] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. A. Riedmiller, “Playing atari with deep reinforcement learning,” *CoRR*, vol. abs/1312.5602, 2013.
- [2] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. P. Lillicrap, K. Simonyan, and D. Hassabis, “Mastering chess and shogi by self-play with a general reinforcement learning algorithm,” *CoRR*, vol. abs/1712.01815, 2017.
- [3] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, 2017.