FIFA Player Analysis

Jonathan Tynan, Zachary Taylor, Conner Smith
{jetynan, ztaylor, csmith95}@stanford.edu

Overview

- **Motivation**: Predicting player improvement and monetary value helps scouts and coaches to understand the factors that make players better and helps soccer clubs to make informed transfers.
- **Models:** We developed a ___________ model to predict the log of players' 2019 market value and whether or not the player's overall rating improved in 2019 based on the FIFA video game datasets' skill-based features, such as speed and stamina, and metadata, such as position and age.
- **Results:** For the player improvement classification task, our best model achieved a test accuracy of 62%.

Related Work

- Dey, Sourya created a multilayer-perceptron neural network to predict the price bracket of players from the FIFA 2017 dataset [1]. The overall accuracy was 40.32%, but 87.2% of predictions were 2 price brackets or fewer away from the true bracket. Additionally, they place any player within 6.32% of his actual price on average. We thought this was a clever way to transform a regression task into a classification task, though in our opinion, a regression task is the more natural expression of the problem.
- Lopez, Waldemar created a support vector regression model to predict the overall rating of players in the FIFA 2018 dataset [2]. They achieved a mean squared error of 1.23 and a variance score of 0.97, which seems to be state of the art in this particular task. Our task differs in that we have different outputs, but many of the features and methods are the same.
- R. Stanojevic and L. Gyarmati used proprietary data from a sports analytics company to predict soccer player market value [3]. The median relative difference between their predictions of market value and actual market value was 34%, and the mean difference was 6% (showing the influence of outliers on the mean error measure. This paper, however, does not consider neural network models.

Data

- **Our dataset consists of 2017, 2018, and 2019 player catalogues from Kaggle.** According to EA Sports, "a network of over 9000 members review the player's abilities, watch him play, and help assign him various ratings" [6]. In each catalogue, every FIFA player is associated with a set of features. We labeled each player in 2019 as improved if their overall rating increased from 2018 to 2019 and not improved otherwise.
- **Features**: **Our dataset contains 36 features total, ranging from skill-based features such as overall rating, speed, stamina, passing accuracy, etc. to metadata such as position, age, and foot preference**.
- **Classification Task Breakdown:**
  - Train set consisted of 8898 players (all from 2017-2018 trajectory) while the dev/test sets consisted of 4449 players, drawn from 2018-2019 trajectory.
  - 3194 players improve in the train set while 1911 and 1929 improved in the dev/test sets. Train/Test means and variances are (0.35, 0.23) and (0.43, 0.25), respectively.
- **Limitations**:
  - Since the players and features vary from year to year, we had to prune from roughly 16 thousand to 9 thousand players.

Figure 1. Logistic Regression performed the best, with a 62% train accuracy.

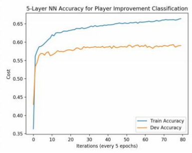

Figure 2. 5-Layer NN overfit slightly. We experimented with Dropout and L2 Regularization to improve this in (Figures 3, 6).

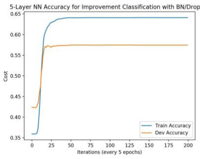

Figure 3. BatchNorm and Dropout had little effect on dev accuracy..

Experiments

Player Improvement Classification Task

- We used scikit to run Logistic Regression, SVM, and RF classifiers. We implemented various NN architectures using TensorFlow. The figures above and to the right show that various hyperparameters/architectures had little effect on classification performance. For all the NN architectures, we used Adam Optimizer.

Player Market Value Regression Task

- We trained a variety of regression models, including naive mean prediction, linear regression, ridge regression, SVM, random forest,, and various neural network architectures.

Results

Player Improvement Classification Task

- All of the models except SVM outperformed the naive baseline model (predicting the mode of the train distribution). Our best model was a Logistic Regression Classifier with a test accuracy of 62%. We found that the best NN architecture was a 5-layer NN (learning rate 0.0001) with BatchNorm, L2 Regularization and Ac

Model	Precision	Recall	Accuracy
Naive	0.00	0.57	0.57
Logistic Regression	0.62	0.63	0.63
SVM	0.33	0.57	0.57
Random Forest	0.59	0.60	0.60
5-layer Neural Network	0.32	0.57	0.57

Player Market Value Regression Task

- For the regression task, our best performing model was a 5 layer neural network with 50 hidden units in each layer. All of our models outperformed the naive baseline. In error analysis, we found that most of the error came from failing to accurately predict the extremely high transfer value of a few young superstars.

Model	Mean Squared Error (after log transformation)
Naive	3.81
Linear Regression	1.95
Ridge Regression	1.95
SVM	1.95
Random Forest	1.69
3-layer Neural Network	1.74
5-layer Neural Network	1.68
>= 7-layer Neural Networks	1.68

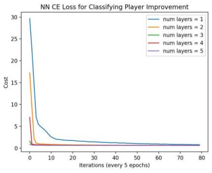

Figure 4. Adding additional NN layers affected convergence time but not the asymptotic cost..

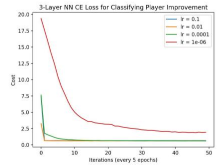

Figure 5. Similar to adding more layers, tweaking the learning rate didn't affect asymptotic training cost.

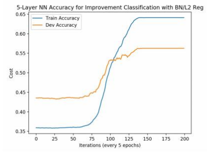

Figure 6. BatchNorm and L2 Regularization didn't significantly bridge the gap between train/dev accuracy.

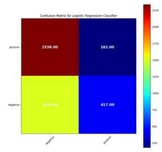

Figure 7. Confusion matrix for our best classifier, logistic regression.

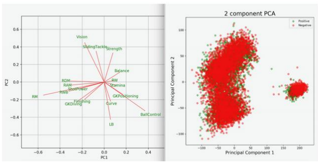

Figure 8. PCA analysis of underlying examples for both categories in the classification task, for the purpose of qualitative analysis. No clear relation to player performance.

Discussion and Future Work

For the player improvement task, it was disheartening to see that the top performer was vanilla logistic regression. For the player market value regression task, we had better results, with a 5-layer NN outperforming all other models.

We believe there's great potential for applying deep learning to sports analytics. One idea is to investigate the PCA plot (section 5.1) further. The clustering suggests some interesting patterns hidden in the data, but our specific project goals failed to uncover/capitalize on these trends. **The PCA clustering clearly doesn't correspond to player improvement, but perhaps this dataset can perform very well on other tasks we didn't explore. Future work should prioritize gathering data that spans more years with the intention of building an RNN that can capture trends over time**. Other projects could analyze data drawn from player statistics (versus FIFA virtual rating system).

References

[1] Dey, Sourya. "Pricing Football Players Using Neural Networks." *ArXiv.org*, 18 Nov. 2017, www.arxiv.org/abs/1711.05865.
[2] Lopez, Waldemar. "Applying Machine Learning to FIFA 2018." *UruIT*, 5 Feb. 2019, www.uruit.com/blog/soccer-and-machine-learning-tutorial/.
[3] R. Stanojevic and L. Gyarmati, "Towards Data-Driven Football Player Assessment," 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), Barcelona, 2016, pp. 167-172.
[4] Gadiya, Karan. "FIFA 19 complete player dataset." Kaggle. 21 Dec. 2018. 17 May 2019 <https://www.kaggle.com/karangadiya/fifa19>.
[5] Shrivastava, Aman. "FIFA 18 Complete Player Dataset." *Kaggle*, 30 Oct. 2017, www.kaggle.com/thec03u5/fifa-18-demo-player-dataset.
[6] Agarwal, Soumitra. "Complete FIFA 2017 Player Dataset (Global)." *Kaggle*, 13 Apr. 2017, www.kaggle.com/artimous/complete-fifa-2017-player-dataset-global.
[7] Saed, Sherif. "EA Explains How FIFA Player Ratings Are Calculated." *VG247*, 27 Sept. 2016, www.vg247.com/2016/09/27/how-ea-calculates-fifa-17-player-ratings/.