Investigation of Overall Rating Fairness in FIFA Games via Machine Learning Methods

Hanze Yao   
*Department of Civil and Environmental Engineering*  
*University of California Berkeley*Berkeley, United States  
yaohanze2016@berkeley.edu

Chao Wang  
*Department of Civil and Environmental Engineering*  
*University of California Berkeley*Berkeley, United States  
chao\_wang@berkeley.edu

***Abstract*—FIFA is a football simulation game published by Electronic Arts. There is a long-existing claim in the FIFA community that English Premier League (EPL) and England players have advantage on overall rating over other players with similar in-game stats. Because England has one of the largest FIFA gaming communities and EPL has the highest market value among all professional football leagues in the world. In this report, classification and neural network methods are used to investigate the overall rating fairness from FIFA 15 to FIFA 20.**

***Keywords—classification, neural network, fairness***

# Introduction

FIFA is an industry-leading football simulation game series developed by EA Vancouver and EA Romania. Each in game player has six categories of stats: pace, dribbling, shooting, defending, passing and physical. Goalkeepers have a different rating system with six stats categories: diving, handling, kicking, reflexes, speed and positioning. Each stats category contains several sub stats. For example, pace includes acceleration and sprint speed. The overall rating model based on in game stats is never open to the public. As a result, FIFA has been criticized for favoring English Premier League and England players on overall rating because of the huge market value of the England FIFA community. An example of this potential overall rating unfairness is as follows. In the latest FIFA game, Ferland Mendy, a French left back playing for Real Madrid in La Liga (top football league in Spain, one of the five major European football leagues), only has 83 overall rating with 2168 total in game stats. While Jamie Vardy, an England striker playing for Leicester City in English Premier League, has 86 overall rating with only 2143 total in game stats. Three overall rating is a big difference for top tier football players in FIFA. In this project classification is used to investigate the overall rating fairness for EPL players. Neural networks are used to investigate the overall rating fairness for England players. The methodology and results are discussed in detail in the following sections.

# Methodology

## Fairness between players from EPL and other leagues

Unlike the debate over COMPAS on the fairness between black and white defendants. In this project, there is not an attribute that can measure the overall rating like the COMPAS score to the risk of recidivism. Therefore, it is not possible to use the same method as setting a threshold for COMPAS score and investigating the false positive rate for different races. So, a new method is used in this project. In order to formulate this as a classification problem, a threshold is required for overall rating. To select a reasonable threshold, only players from five major European Leagues (English Premier Leagues, Spanish La Liga, Germany Bundesliga, Italian Serie A and French Ligue 1) are considered. Because choosing a single overall rating threshold can guarantee a similar percentage of players with overall above and below the threshold between different leagues. The mean and median overall for the five leagues are shown in Table 1. A threshold of 73 is selected for classification.

1. Mean and Median Overall for Different Leagues

|  |  |  |
| --- | --- | --- |
| **Game** | **English Premier League** | **Other Four European Leagues** |
| FIFA 15 | 70.5 / 72.0 | 69.9 / 71.0 |
| FIFA 16 | 72.9 / 75.0 | 71.9 / 73.0 |
| FIFA 17 | 72.2 / 75.0 | 72.6 / 74.0 |
| FIFA 18 | 72.5 / 75.0 | 72.5 / 73.0 |
| FIFA 19 | 72.6 / 75.0 | 72.8 / 74.0 |
| FIFA 20 | 73.4 / 75.0 | 72.4 / 73.0 |

Now this is a binary classification problem for players with 73 overall or above and players with below 73 overall based on 8 features: pace, dribbling, shooting, defending, passing, physical, skill move and weak foot. Goalkeepers are not considered because of small data size. SGDClassifier from sklearn.linear\_model with l2 regulation and logistic loss is used. After training, calculate the False Negative Rate (FNR) and False Positive Rate (FPR) of the test set. Now apply the EPL classification model to the test set of four other major European Leagues. If there is an overall rating boost for EPL players, the FNR should decrease and the FPR should increase compared to the original FNR and FPR of the four other major European Leagues’ test set.

## Fairness between players from England and other countries

In this subsection, the fairness between England football players and football players from other countries is studied. A three layer fully connected neural network is trained to predict the overall rate of a player based on his stats attributes. The number of nodes for the two hidden layers of this neural network are 25 and 36 respectively. The active functions are chosen as the rectified linear unit (ReLU).

From FIFA 15 to FIFA 20, in each year, two different models are then built with the same architecture given above. The difference is the training data sets. One model is trained with England players, the other model is trained with all non-England players. Finally, the model trained with England players is used to predict the overall rating of the non-England players. If for non-England players, the overall rate predicted by the model trained with England players is higher than the value predicted by the model trained with England players, we then can conclude the bias in the FIFA rating model.

# Results

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections A-D below for more information on proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

## Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

## Units

* Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as “3.5-inch disk drive”.
* Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.
* Do not mix complete spellings and abbreviations of units: “Wb/m2” or “webers per square meter”, not “webers/m2”. Spell out units when they appear in text: “. . . a few henries”, not “. . . a few H”.
* Use a zero before decimal points: “0.25”, not “.25”. Use “cm3”, not “cc”. (*bullet list*)

## Equations

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

Number equations consecutively. Equation numbers, within parentheses, are to position flush right, as in (1), using a right tab stop. To make your equations more compact, you may use the solidus ( / ), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in:

*a**b* 

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the equation. Use “(1)”, not “Eq. (1)” or “equation (1)”, except at the beginning of a sentence: “Equation (1) is . . .”

## Some Common Mistakes

* The word “data” is plural, not singular.
* The subscript for the permeability of vacuum *μ*0, and other common scientific constants, is zero with subscript formatting, not a lowercase letter “o”.
* In American English, commas, semicolons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
* A graph within a graph is an “inset”, not an “insert”. The word alternatively is preferred to the word “alternately” (unless you really mean something that alternates).
* Do not use the word “essentially” to mean “approximately” or “effectively”.
* In your paper title, if the words “that uses” can accurately replace the word “using”, capitalize the “u”; if not, keep using lower-cased.
* Be aware of the different meanings of the homophones “affect” and “effect”, “complement” and “compliment”, “discreet” and “discrete”, “principal” and “principle”.
* Do not confuse “imply” and “infer”.
* The prefix “non” is not a word; it should be joined to the word it modifies, usually without a hyphen.
* There is no period after the “et” in the Latin abbreviation “et al.”.
* The abbreviation “i.e.” means “that is”, and the abbreviation “e.g.” means “for example”.

An excellent style manual for science writers is [7].

# Discussion

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

## Authors and Affiliations

**The template is designed for, but not limited to, six authors.** A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

### For papers with more than six authors: Add author names horizontally, moving to a third row if needed for more than 8 authors.

### For papers with less than six authors: To change the default, adjust the template as follows.

#### Selection: Highlight all author and affiliation lines.

#### Change number of columns: Select the Columns icon from the MS Word Standard toolbar and then select the correct number of columns from the selection palette.

#### Deletion: Delete the author and affiliation lines for the extra authors.

## Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is “Heading 5”. Use “figure caption” for your Figure captions, and “table head” for your table title. Run-in heads, such as “Abstract”, will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced. Styles named “Heading 1”, “Heading 2”, “Heading 3”, and “Heading 4” are prescribed.

## Figures and Tables

#### Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

1. Table Type Styles

|  |  |  |  |
| --- | --- | --- | --- |
| **Table Head** | **Table Column Head** | | |
| ***Table column subhead*** | ***Subhead*** | ***Subhead*** |
| copy | More table copya |  |  |

1. Sample of a Table footnote. (*Table footnote*)
2. Example of a figure caption. (*figure caption*)

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

# Conclusion

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

##### Acknowledgment *(Heading 5)*

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

##### References

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more, give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

1. G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*
2. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
3. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
4. K. Elissa, “Title of paper if known,” unpublished.
5. R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
6. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
7. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**

