HW3: Generalized Linear Models

Žan Pušenjak
MLDS 2024/25 , FRI, UL
zp4613@student.uni-lj.si

***Abstract*—We implemented multinomial and ordinal logistic regression models and tested them on a provided data set of basketball shot to observe the performance of both models and provide insights into their training.**

I. Implementation

We implemented both models by minimizing their log-likelihood with the help of the *scipy* function $fmin\_l\_bfgs\_b$[1].

To achieve accurate interpretability, we used the first class as *reference class*. With that we achieved the right number of parameters to optimize:

- **Multinomial** $(M-1)K$
- **Ordinal** $(M-2)+K$

Where M is the number of classes and K is the number of features.

II. Application of the Multinomial Regression

We used the provided data set of basketball shot types[2] and our model implementation to provide some insight into the relationship between the *ShotType* feature and other features.

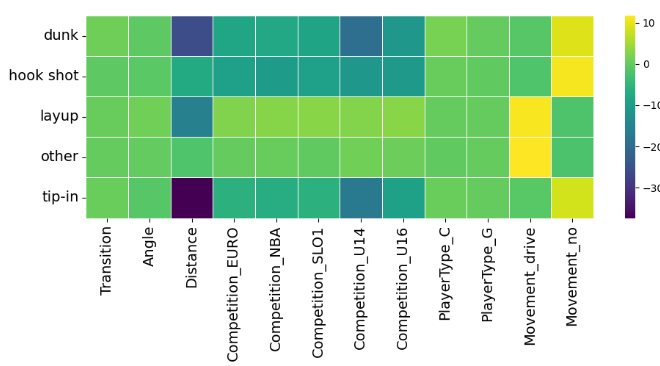

Fig. 1. The values of beta parameters on the provided dataset

We additionally prepared the data by one hot encoding the categorical features *Competition*, *PlayerType* and *Movement* and normalizing the numerical features *Distance* and *Angle*. Finally we removed highly correlated features (features with correlation coefficient > 0.7), which allowed us to better interpret the final parameters.

We used the β parameter matrix as the main instrument for our findings. A β parameter represents the **log-odds** of observing the class as opposed to observing the reference class. So a higher β value will increase those odds for the class and a lower value will decrease them. Our **reference class** was *above-head*

Observing Figure 1 we selected the highest and lowest β and tried to find any explainable relations.

A. Competition U14

We can observe that two β parameters for the *Competition_U14* feature are negative, values being:

- *dunk*: $\beta = -19.88 \pm 3.98$
- *tip-in*: $\beta = -17.32 \pm 3.41$

So if the shot was recorded in a U14 basketball game the chances of observing a *dunk* or a *tip-in* rather than an *above-head* shot are significantly decreased. This is a logical explanation since the players at that age are usually not tall enough to perform a shot of that type and mostly throw *above-head*.

B. Movement

The *Movement_drive* and *Movement_no* features positively influence the *hook-shot*, *layup* and *other* types, with the exact parameter values being:

1) for *Movement_drive*
   - *layup*: $\beta = 11.27 \pm 1.67$
   - *other*: $\beta = 11.70 \pm 1.66$
2) for *Movement_no*
   - *dunk*: $\beta = 9.06 \pm 1.94$
   - *hook-shot*: $\beta = 11.02 \pm 2.33$
   - *tip-in*: $\beta = 8.31 \pm 1.67$

This tells us that the chances of observing a *layup* or *other* shot type instead of *above-head* will increase if there is *drive* movement present. The unexpected results that we could not interpret were that if there is no movement present a *dunk*, *hook-shot* or *tip-in* will occur more regularly.

On the other hand

[1] https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.fmin_l_bfgs_b.html
[2] https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0128885

C. Distance

It is apparent the *Distance* feature was the most important feature on some classes. This is expected, as the further the player is from the rim, the less likely he is to perform a *tip-in*, *dunk* or *layup*, in that order, and that reflects on the β parameters.

- *tip-in*: $\beta = -37.34 \pm 3.86$
- *dunk*: $\beta = -25.97 \pm 2.97$
- *layup*: $\beta = -16.12 \pm 1.15$

III. Application of the Ordinal Regression

We made some synthetic data using a data generating process (DGP) of ordinal value to show that the ordinal regression model will outperform the normal multinomial regression with suck DGP. We used accuracy as the metric for comparison. The performance of the models can be observed on Figure 2 where we observe that the ordinal regression clearly outperforms the multinomial one. We can also see that increasing the data set does not attribute to the results.

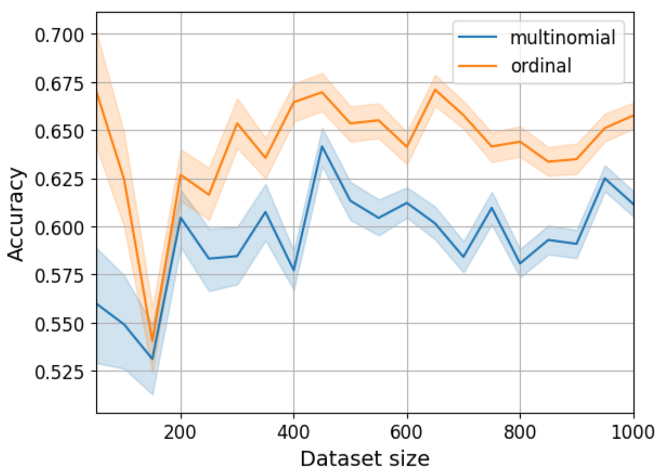

Fig. 2. Performance on an ordinal dataset at different data set sizes.

Even when increasing the data set size to 10000 the ordinal regression still achieved a better accuracy seen in Table I.

TABLE I
Accuracy of ordinal vs multinomial regression with ordinal data set of size 10000.

Model	Accuracy
Ordinal reg.	0.59 ± 0.01
Miltinomial reg.	0.64 ± 0.01