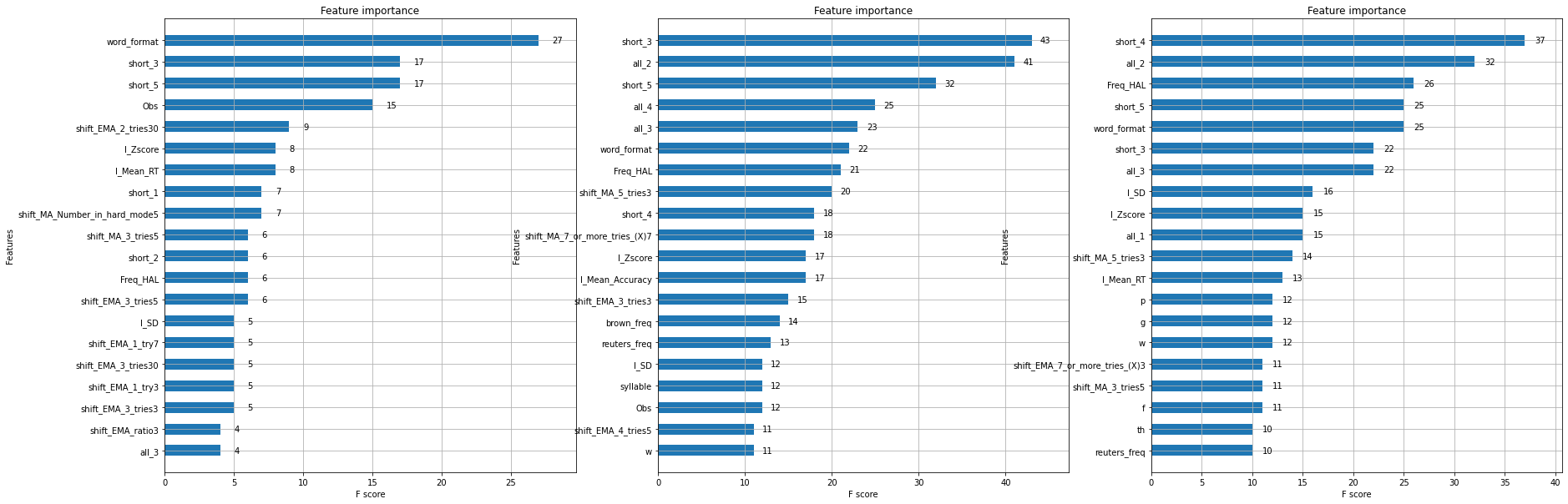
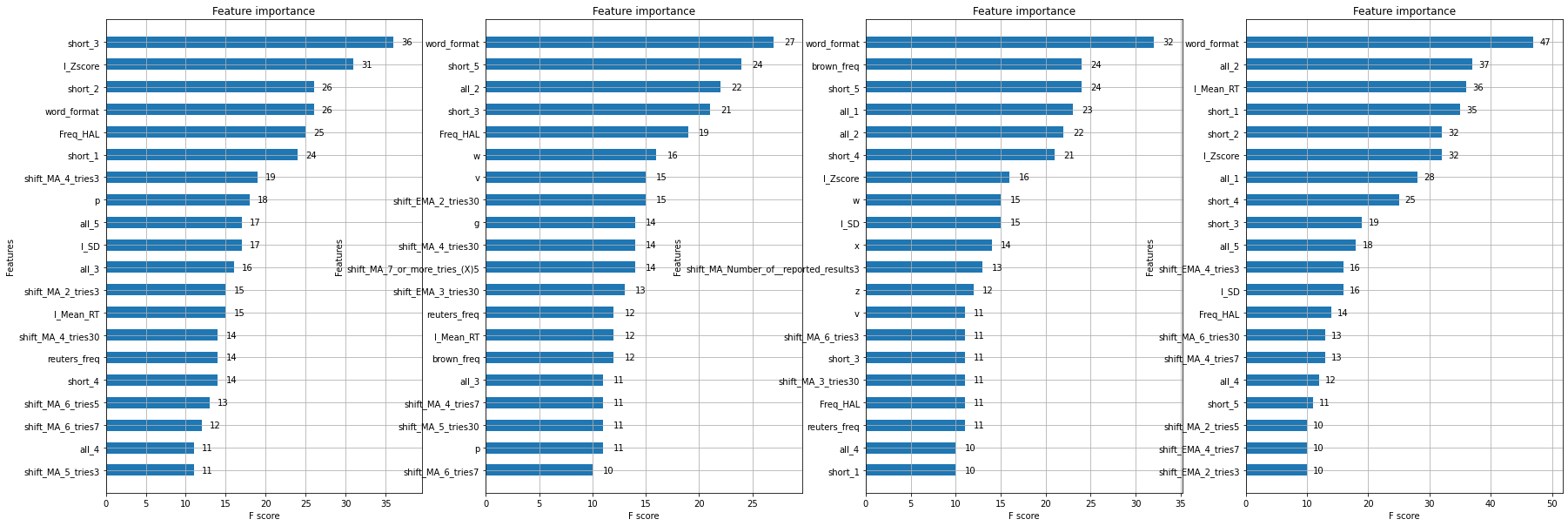
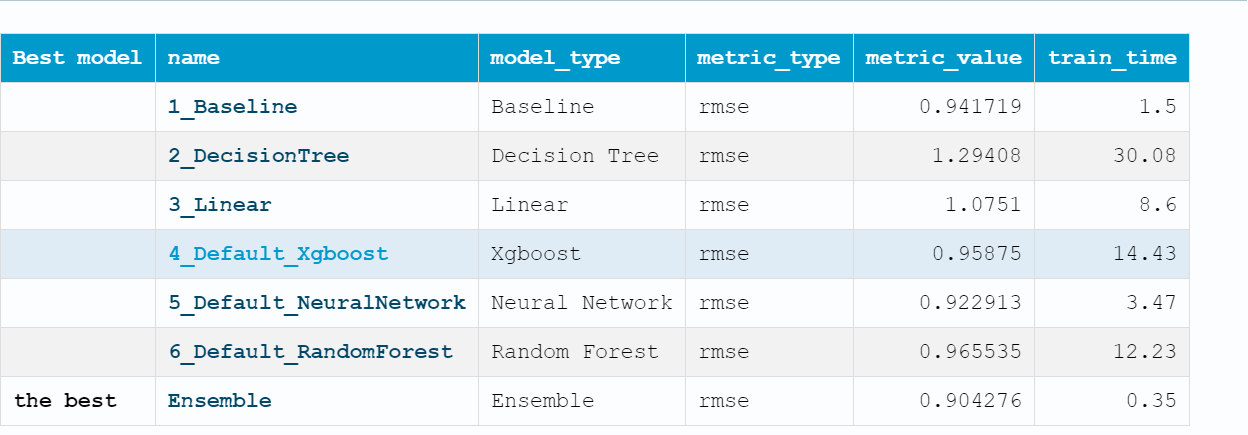
In this problem, we try to combine times series with word attributes. For time series, we take the moving average and weighted moving average 3,5,7,30 days. Since the date, we will predict is the 03/01/2023, the latest data we have is on 12/31/2022. There are 60 days in between. We also make our training data to predict the distribution of tries 60 days later. Therefore, our weighted average needs also are shifted to 60 days. This means that our prediction will use the moving average from 60 days before not its current moving average. We also add the day of the week information to the dataset. With all these time series and word attributes information, we also use XGboost to predict our distribution.

The importance of features for 1,2,...,7 tries are as follows:







The result of the distribution of the prediction is 0.323888, 3.1039689, 10.723816, 17.66941, 19.789255, 11.09584, 5.549209