

HW3: Error Analysis

Submitted By: Mayanka Jha

The followings are some of the rows from the test dataset whose values have been predicted incorrectly.

Row	Correct Value	Predicted Value
32	FALSE	TRUE
33	FALSE	TRUE
88	FALSE	TRUE
91	FALSE	TRUE

Using below code, we can see some of the incorrectly predicted label values:

```
print(test['body_score'].values[index])
print(test['tournaments'].values[index])
print(test['answer_type'].values[index])
print(test['inlinks'].values[index])
print(test['tournament_year'].values[index])
print(test['length'].values[index])
```

Row 32: The protagonist of this novella is inspired to travel by the sight of a man standing before a mortuary chapel, while symbolically important is his encounter with an old man who has joined a group of young clerks by donning a wig and false teeth.

27.5387991702

2001 ACF Fall

work

1

2001

45

Row 33: The protagonist of this novella is inspired to travel by the sight of a man standing before a mortuary chapel, while symbolically important is his encounter with an old man who has joined a group of young clerks by donning a wig and false teeth.

26.976121095

2001 ACF Fall

work

2

2001

45

Row 88: In a symbolic incident in this novel Ronny Heaslop and his friend are unsuccessful in their attempts to organize a bridge game.

45.8488311908

2001 ACF Regionals

work

5

2001

22

Row 91: We first meet Fiers, the aging footman, in Act I and learn that he is to be replaced by Yasha. Other minor characters include Charlotta, a governess, and Gaev, the brother of the central female. Near the end of the play, one character rejoices in his new possession exclaiming, "If my father and grandfather could see me now - me, their ignorant Yermolay, almost illiterate." Lubov Andreyevna, aka Madame Ranevsky, faints when she is forced to part with the dearest reminder of her dead son to Lopakhin, a former family friend and newly wealthy man.

99.8111686167

2001 ACF Regionals

work

11

2001

96

The wrong values seem to be predicted for rows having tournament year 2001, answer type work and varying lengths of text. Our model is not robust enough using the features that we have chosen. For future work we can create a new body score to improve our prediction accuracy.

