

Text classification on comments of Genshin Impact via non-deep learning and deep learning models

732A92 Text Mining

Chenjian Shi (chesh532)

Jan 15,2022

Abstract

test

Contents

1	Introduction	1
2	Data	1
2.1	Data Access	2
2.2	Data Overview	2
2.3	Data Cleaning	2
3	Process	2
3.1	Non-Deep Learning	2
3.2	Deep Learning	2
4	Conclusion	2
5	Reference	2

1 Introduction

2 Data

```
##          Time  Score  \
## 0    2022-01-05 04:22:08    5
## 1    2022-01-07 04:54:02    5
## 2    2022-01-08 00:08:37    5
## 3    2022-01-04 23:52:11    5
## 4    2022-01-05 01:08:32    5
## ...          ...    ...
## 38593 2021-10-14 09:23:25    1
## 38594 2021-09-29 11:21:58    1
## 38595 2021-09-29 05:19:42    1
## 38596 2021-09-28 20:33:23    1
## 38597 2021-09-29 11:05:05    1
##
##
```

Content

```

## 0      I literally love this game. Playing on a phone...
## 1      The game itself is wonderful and enchanting, I...
## 2      This is an amazing game it has amazing graphic...
## 3      Full fledged game that you can play for free. ...
## 4      This game is so fun! The story is thrilling an...
## ...
## 38593  I can't even play it.I just dowloaded it and I...
## 38594  The game has been fun so far, although the upd...
## 38595  Aniversary awards sucks as hell. This game dev...
## 38596  Overall a very good game with very good story....
## 38597  This is the only game in my life I've played f...
##
## [38598 rows x 3 columns]

```

2.1 Data Access

2.2 Data Overview

2.3 Data Cleaning

3 Process

3.1 Non-Deep Learning

3.1.1 Data preprocessing

3.1.2 LSVC

3.1.3 Logistic Regression

3.2 Deep Learning

3.2.1 Data Preprocessing

3.2.2 RNN

3.2.3 LSTM

3.2.4 BERT

4 Conclusion

5 Reference