



Text Classification with Multiple-Choice Model Website

Prepared by

Abdel-Rhman Sayed EL-Sadory

Water Engineer

*Special Thanks for the most awesome and
professional engineers*

Engineer-**Sayed Omar**
Engineer-**Abdo Gharib**

Engineer-**Shady Nagy**
Engineer- **Doaa Mohamed**

Agenda

- Who am I
- Why Data Science !?
- Problem Type & The Process
- Technical Details
 - Dataset
 - Discovering Data
 - Model Selection and Evaluation
- The Website
- The Website (Tricks)

Who am I

- Graduated from faculty of science (Petroleum Geology Dep.)
- Water Engineer (Treatment)
 - Al Sharqia Sugar Manufacturing S.A.E (Al-Nouran)
 - Veolia Water Technologies

Why Data Science !?

- Passion
- Career Shifting



Problem Type & The Process



- De-Noising & Removing Stop-words



- Text Converting & Label Encoding



- Building the Models



- Create The API



- Link The Processing Code With The Website

Technical Details (Dataset)

- Dataset from **IMDb** website.


	A	B
1	review	sentiment
2	Kind of drawn in by the erotic scenes, only to realize this was one of the most ama	negative
3	One of the other reviewers has mentioned that after watching just 1 Oz episode yc	positive
4	A wonderful little production. The filming technique is very unassumin	positive
5	I thought this was a wonderful way to spend time on a too hot summer weekend, s	positive
6	Basically there's a family where a little boy (Jake) thinks there's a zombie in his clo	negative
7	Petter Mattei's "Love in the Time of Money" is a visually stunning film to watch. Mi	positive
8	Probably my all-time favorite movie, a story of selflessness, sacrifice and dedicatio	positive
9	I sure would like to see a resurrection of a up dated Seahunt series with the tech th	positive
10	This show was an amazing, fresh & innovative idea in the 70's when it first aired. Th	negative
11	Encouraged by the positive comments about this film on here I was looking forward	negative
12	If you like original gut wrenching laughter you will like this movie. If you are young	positive
13	Phil the Alien is one of those quirky films where the humour is based around the o	negative
14	I saw this movie when I was about 12 when it came out. I recall the scariest scene v	negative
15	So im not a big fan of Boll's work but then again not many are. I enjoyed his movie	negative
16	The cast played Shakespeare. Shakespeare lost. I appreciate	negative
17	This a fantastic movie of three prisoners who become famous. One of the actors is	positive
18	Some films just simply should not be remade. This is one of them. In and of itself it	positive
19	This movie made it into one of my top 10 most awful movies. Horrible. 	negative
20	I remember this film, it was the first film i had watched at the cinema the picture w	positive
21	An awful film! It must have been up against some real stinkers to be nominated fo	negative
22	After the success of Die Hard and it's sequels it's no surprise really that in the 1990s	positive
23	I had the terrible misfortune of having to view this "b-movie" in it's entirety. 	negative

Technical Details (Discovering Data)

- No Nulls

```
IM.data.info()


<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2000 entries, 0 to 1999
Data columns (total 2 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   review      2000 non-null    object
1   sentiment    2000 non-null    object
dtypes: object(2)
memory usage: 31.4+ KB
```



- Well-Distributed Data

```
IM.data["sentiment"].value_counts()

positive    1005
negative     995
Name: sentiment, dtype: int64
```



Technical Details (Discovering Data)

- After Data Cleaning

	review	sentiment	Cleaned_Reviews
0	Kind of drawn in by the erotic scenes, only to...	negative	kind drawn erotic scenes realize one amateurs...
1	One of the other reviewers has mentioned that ...	positive	one reviewers mentioned watching 1 oz episode ...
2	A wonderful little production.

 The...	positive	wonderful little production br br filming tech...
3	I thought this was a wonderful way to spend ti...	positive	thought wonderful way spend time hot summer we...
4	Basically there's a family where a little boy ...	negative	basically there's family little boy jake think...

Technical Details (Model Selection and Evaluation)

Model	Evaluation (Dataset 1000)	Evaluation (Dataset 2000)
Logistic Regression	0.81	0.80
KNN	0.52	0.54
Naïve Bayes	0.57	0.65
SVM (Linear)	0.795	0.79
SVM (Poly)	0.58	0.57
SVM (rbf)	0.775	0.777

- Logistic Regression & SVM(rbf) are the best.
- Although the SVM(linear) Evaluation is 0.79, but the prediction is bad and not practical.

The Website

The Website (Tricks)

- Select Tag
- Iframe Tag

Thanks !