

Klasifikácia YouTube komentárov pomocou neurónovej siete

MAROŠ KOPEČ

Masarykova Univerzita

487595@muni.cz

June 20, 2019

I. ÚVOD

Klasifikácia spamu je problémom, ktorému je venovaných množstvo odborných článkov a prác. Za spam sa považuje nevyžiadaná správa rozosielená veľkému počtu adresátov alebo na mnoho miest, zväčša za účelom reklamy. Pre množstvo internetových stránok je práve spam veľkým problémom, pretože ich používatelia sú zahltený obsahom, ktorý im znepríjemňuje skúsenosť s ich obsahom. Práve YouTube je terčom veľkého množstva spamu od používateľov, ktorí sa týmto spôsobom snažia zviditeľniť vlastné video či kanál.

II. SÚVISIACE PRÁCE

Táto práca je inšpirovaná prácou TubeSpam¹ z Federálnej Univerzity Sao Carlos. V tejto práci porovnávali niekoľko metód klasifikácie spamu, menovite rozhodovacie stromy, K-najbližších susedov, logistická regresia, Bernoulliho naivný Bayes, Gaussov naivný Bayes, Multinomiálny naivný Bayes, náhodné lesy, Support vector machines s lineárnym kernelom, Support vector machines s polynomiálnym kernelom a Support vector machines s Gaussovým kernelom. Práca dopĺňa vyššie zmienenú o modely postavené na rekurentných neurónových sieťach.

¹Alberto, T.C., Lochter J.V., Almeida, T.A. TubeSpam: Comment Spam Filtering on YouTube. Proceedings of the 14th IEEE International Conference on Machine Learning and Applications (ICMLA'15), 1-6, Miami, FL, USA, December, 2015. (preprint)

Table 1: Kompozícia datasetu

Dataset	YouTube ID	# Spam	# Ham	Total
Psy	9bZkp7q19f0	175	175	350
KatyPerry	CevxZvSjLk8	175	175	350
LMFAO	KQ6zr6kCPj8	236	202	438
Eminem	uelHwf8o7_U	245	203	448
Shakira	pRpeEdMmmQ0	174	196	370

V pôvodnej práci klasifikovali spam oddelene pre každé video. Neurónová sieť potrebuje veľké množstvo záznamov aby bola schopná naučiť sa rozpoznávať spam. Dataset však obsahuje len obmedzený počet záznamov. Preto sa dataset spojil a všetky experimenty sa vykonávali nad zmiešanými záznamami.

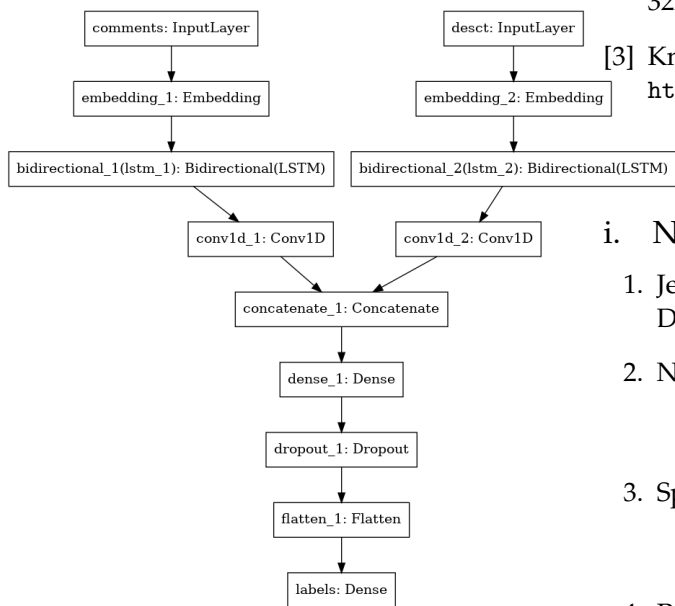
III. DATASET

Pre trénovanie a testovanie modelov sú použité rovnaké dáta ako boli použité v práci TubeSpam². Dataset obsahuje 1956 záznamov z 5 najpozeranejších YouTube videí vo formáte csv. Kompozícia datasetu je znázornená v tabuľke ???. Pre potreby experimentu bol dataset rozšírený o popis videa.

Pre predstavu obsahu záznamov boli vygenerované slovné mapy, ktoré sú zobrazené na obrázkoch ??? a ???. Nad záznamami bol vykonaný experiment, pri ktorom bolo náhodne vybraných 100 záznamov z videa od autora Eminem, ktoré boli znovu označované. Z týchto sto záznamov bolo 8.9% označených

²<https://archive.ics.uci.edu/ml/datasets/YouTube+Spam+Collection>

Figure 4: Model č.2



V. VÝSLEDKY

Výsledky k prvému navrhnutému modelu ?? so vstupom len pre komentáre sa žiaľ nepodarilo kvôli technickej chybe zachovať. K dispozícii sú len dáta z optimalizácie hyperparametrov. Tieto výsledky boli zohľadnené pri experimentoch s druhým modelom ??.

Metódy evaluácie Pre zhodnotenie modelu boli použité štatistické metriky *presnosť*, *chytený spam*, *blokovaný ham*, *F-measure*, *Matthews korrelačný koeficient*, ktoré boli použité aj pri evaluácii v spomínanej práci TubeSpam.

REFERENCES

- [1] Michel Goossens, Frank Mittelbach, and Alexander Samarin. *The L^AT_EX Companion*. Addison-Wesley, Reading, Massachusetts, 1993.
- [2] Albert Einstein. *Zur Elektrodynamik bewegter Körper*. (German) [*On the electrodynamics*

of moving bodies]. *Annalen der Physik*, 322(10):891–921, 1905.

- [3] Knuth: *Computers and Typesetting*, <http://www-cs-faculty.stanford.edu/~uno/abcde.html>

VI. PRÍLOHA

i. Návod na spustenie

1. Je potrebné stiahnuť vstupné dáta. Dataset ⁴, Embedding ⁵
2. Nainštalujte si nástroj:
pipenv
3. Spustite príkazy:
pipenv install
4. Pre spustenie tréovania modelu použite príkaz:
python classifier.py NUMBER

Kde NUMBER je číslo experimentu a meno zložky výstupov.

⁴<https://archive.ics.uci.edu/ml/machine-learning-databases/00380/>

⁵<https://nlp.stanford.edu/projects/glove/>