Task 1: basic methods in text classification
School of Computer Science and Electronic Engineering
University of Essex
CE807-7-SU: Text Analytics
registration number:2110366

basic methods in text classification

Machine Learning is another field of AI learning that does not define rules. It is about telling or suggesting information that AI can know and understand by itself[2]. Machine learning also has techniques related to text classification varioussuch as SVM, Logistics Regression, etc[3],[1].

The Logistic Regression or Maximum Entropy Model (MaxEnt) is the most important machine learning model. because it does not require a lot of tuning and can be developed to be applied quickly and fairly effective without much effort

Machine learning has limitations in text classification. Artificial intelligence classification is based mainly on 'data', large volumes of data and accurate data. If the data is biased or is biased then artificial intelligence will bias as well, even if it is right data or the more the AI picks up the wrong data. It will also misinterpret that information[4].

REFERENCES

[1] Atreya Basu, Carolyn Watters, and Michael Author. 2003. *Support Vector Machines for Text Categorization*. 103.
   DOI: <http://dx.doi.org/10.1109/HICSS.2003.1174243>

[2] M Ikonomakis, Sotiris Kotsiantis, and V Tampakas. 2005. *Text classification using machine learning techniques. WSEAS transactions on computers* 4, 8 (2005), 966–974.

[3] Tomas Pranckevicius and Virginijus Marcinkevic̆ius. 2017.
   Comparison of Naive Bayes, Random Forest, Decision Tree, Support Vector Machines, and Logistic Regression Classifiers for Text Reviews Classification. Baltic Journal of Modern Computing 5 (01 2017).
   DOI: <http://dx.doi.org/10.22364/bjmc.2017.5.2.05>

[4] Kanish Shah, Devanshi Sanghvi, and Manan Shah. 2020.
   A Comparative Analysis of Logistic Regression, Random Forest and KNN Models for the Text Classification. Augmented Human Research 5 (12 2020).
   DOI:<http://dx.doi.org/10.1007/s41133- 020- 00032- 0>