Sentiment Analysis for Video Libraries

# **Introduction**

Computational intelligence methods are proving to be valuable strategic tools in many industries with the increasing popularity of analytics and data science (Yang, Yin and Mo, 2015). For instance, data is mined for patterns in business analytics that would help better understand customers and improve sales and marketing. Methods of computational intelligence make it possible to use probabilistic methods to find patterns in data. Typically, these methods work on low-level data and are not guided by absolute knowledge, as with general artificial Intelligence methods. In addition, a huge amount of data is now generated in written form which warrants analysis. The written text is subject to interpretation, and it is difficult to represent the data in an absolute syntax (such as a binary system). Computational intelligence methods, however, require such fluidity and may be the most appropriate methods to find trends in such data, hence sentiment analysis comes in.

According to Oxford Sentiment analysis is the process of computationally identifying and categorizing opinions expressed in a piece of text, especially in order to determine whether the writer's attitude towards a particular topic, product, etc. is negative, positive or neutral. This subject brings together various research areas such as natural language processing, data mining and text mining, and is rapidly becoming of major importance to organizations as they strive to integrate computational intelligence methods into their operations and attempt to shed more light on their products and services and improve them. Sentiments are opinions or views about a certain object. This concept can be applied in different areas that includes marketing, social media monitoring, brand monitoring, customer feedback, customer service.

Sentiment analysis can be applied in libraries. A library is a collection of resources in a variety of formats that is (1) organized by information professionals or other experts who (2) provide convenient physical, digital, bibliographic, or intellectual access and (3) offer targeted services and programs (4) with the mission of educating, informing, or entertaining a variety of audiences (5) and the goal of stimulating individual learning and advancing society as a whole (Eberhart, 2010). Libraries are key elements of a healthy community so they are vital in a society. To maintain their vitality to the society they need to be maintained based on customers’ view. Thereby incorporating sentiment analysis will help in easily visualising the view of customers on the resources in a library.

# **Background**

Sentiment analysis is a new computational intelligence subject that has a number of researches in different fields that ranges from management science to computer science, social science and business (D’Andrea *et al.*, 2015). A number of researches were done with the aim of trying to apply this concept to different areas.

Sentiment analysis has been applied in hotel systems (Yang, Yin and Mo, 2015). Yang and Yin used sentiments from TripAdvisor with the aim of proving that machine learning techniques of analysing sentiments are better than human produced sentiments. The output proved that machine language is faster, thus it saves time and saves cost to hotels.

In the entertainment industries sentiments researches where done on its application to platform such as YouTube. (Asghar *et al.*, 2015) did a research on the application of Sentiment analysis to YouTube. A number of challenges where met that include failure to analyse different languages that YouTube customer comments with.

(Bogicevic *et al.*, 2013) also did a research about application of sentiment analysis in airports. The research was mainly targeted to see the view of customers on the services offered by the airport and see which services are distractors and which one are enhancers of customers travel on the airport. This is application of sentiment analysis in the field of customer satisfaction.

Customer involvement in the day to day operation of any firm is of uttermost important and sentiment analysis makes this process easier. Sentiment analysis let the voice of customers be heard out. So it brings advantages to the customers of letting them complains be heard and at the same time it helps organisations to maintain their market share since customers will remain loyal to an organisation that listen to their objectives. Sentient analysis brings a market advantage to affirm that uses it.

The application of sentiment analysis to libraries is also an avenue that remains unexploited. With the introduction of Library 2.0 there is need for interacting with the customers to see their view to information contained in the library and help the library owner with decision making.

# **Problem Statement**

The web has created jobs for a number of people through video sharing, article writing, and music uploading, this calls the need for library owners to understand what viewers of the resources in the libraries are saying so as to maintain market share. Analysing large volumes of comments left by the resource viewers is tedious, time consuming and sometimes consuming, so there is no system that helps with viewing the comments and visualise them for the library owner.

# **Aim**

Design a model that does sentiments mining from library comments.

# **Objectives**

* To train the Support Vector Machine.
* To retrieve textual feedback from the customers.
* To omit noisy feedback / data during analysis.
* To classify the feedback from customers into positive and negative, to help in feedback review using Support Vector Machine Algorithm.
* To generate periodic reports of analysed information and visualize it using bar graph.

# **Justification**

There is a number of valid reasons why this system must be produced. The reasons include allowing customers to comment to a resource, reduction of task overload on the library owner and alert on customer dissatisfaction on library resources.

Firstly, the sentiment analysis system for libraries helps in evaluating the extent to which customers are satisfied by the resources offered in a library. This is achieved through the use of a commenting section where customers drop their comments on the resources they are receiving expressing their view and feelings.

In addition, the system will help libraries owners by alerting them about customers’ dissatisfaction before it is worse. The system will visualize the sentiments reports of every month so that the library owner sees the satisfaction rate of customers. Basing on the reports the library owner decides the necessary actions that will help maintain the market share or keep the customers satisfied on the next resource that will be uploaded.

Lastly, sentiment analysis saves time of going through each and every comment that customers leave one by one. The system will analyse the comments and produce a report that summarise the comments from patrons instead of library owner analysing the comments on their own.

# **Methodology**

In this project Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology is going to be adopted. CRISP DM is a methodology that is applied much in data mining but it is considered to be a standard methodology applied to the extraction of knowledge from big data sets (Sharda, Delen and Turban, 2016). Sentiment analysis works with datasets for training of systems either training the system using supervised learning algorithms or unsupervised learning algorithm. Not every step that the CRISP-DM is going to be used, so the methodology will be twisted a bit and neglect some steps.

# **Scope**

The project will focus on the development of a sentiment analysing system for video resources in libraries.

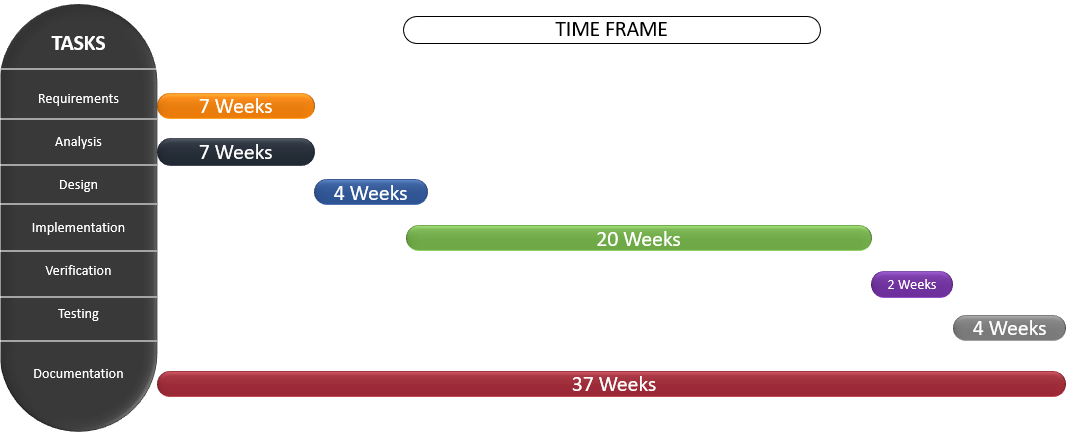
# **Expected Outcome**

After the project is done, the system must be allowing resource viewers to comment and store the comments in a database. This allows library owners to see the view of their customers on the resources they are providing and measure their validity and customer satisfaction.

Moreover, the system must produce periodic reports that visualise customer satisfaction rate to the library owner. For example, the library owner can request a report from 1 June 2019 to 31 December 2019, to see the performance of the library for the last half of the year. This saves the library owner’s time and effort as there is no need to analyse the comments in person. The information will be used for decision making such a s to continue providing the customer the same content or not.

Lastly the system must omit noisy data/ feedback. Noisy data is nonsensical data that if even combined or interpreted won't produce any useful information. Analysis of that data will be just wasting of time so omitting it will be the best thing. Omission of some data will increase the speed of the system.

# **Project Schedules**



# **References**

Asghar, M. Z. *et al.* (2015) ‘Sentiment Analysis on YouTube: A Brief Survey’, pp. 1–11. Available at: http://arxiv.org/abs/1511.09142.

Bogicevic, V. *et al.* (2013) ‘Airport service quality drivers of passenger satisfaction’, *Tourism Review*. Edited by D. Fevzi Okumus, Dr Robertico Croes, D. Emerald Group Publishing Limited, 68(4), pp. 3–18. doi: 10.1108/TR-09-2013-0047.

D’Andrea, A. *et al.* (2015) ‘Approaches, Tools and Applications for Sentiment Analysis Implementation’, *International Journal of Computer Applications*, 125(3), pp. 26–33. doi: 10.5120/ijca2015905866.

Eberhart, G. M. (2010) *The Librarian’s Book of Lists: A Librarian’s Guide to Helping Job Seekers - Google Books*. Chicago: American Library Association. Available at: https://books.google.co.zw/books?hl=en&lr=&id=STr95tKBvZsC&oi=fnd&pg=PP1&dq=The+Librarian’s+Book+of+Lists&ots=uSvL2-7Bjy&sig=UcQsTndmRDQevH5cvhH6RUIGd8c&redir\_esc=y#v=onepage&q=The Librarian’s Book of Lists&f=false (Accessed: 13 November 2019).

Yang, S. Y., Yin, S. and Mo, K. (2015) ‘Social Media and News Sentiment Analysis for Advanced Investment Strategies Behavioral finance and its investment applications View project Social media influencers and contagion modeling View project Social media and news sentiment analysis for advanced ’, (January). doi: 10.1007/978-3-319-30319-2.