

# Data Analysis on E-learning Courses Delivered Online

Rajesh karumanchi, R.Rajkumar

School of Computer Science and Engineering, Vellore Institute of Technology

## ABSTRACT

European and Asian nations, of late, have seen an ascent in the quantity of understudies utilizing the Internet for E-learning as it encourages move of information through coordinated just as non-concurrent techniques. Courses accessible online assistance them in picking up aptitude in the space of their advantage. We need generous measures to assess the viability of these courses. The majority of the e-learning online interfaces contain a gathering to give input through remarks and appraisals. In this paper, we arrange these remarks and appraisals as sure, negative and nonpartisan by utilizing Support Vector Machine (SVM). Disposing of manual boundary tuning for text grouping is best since it lessens the general time taken. SVM is utilized as it can discover great boundary settings naturally. The aftereffect of this investigation helps in evaluating the viability of e-learning courses gave by European and Asian Universities.

## Keywords

E-Learning, Text classification, Sentiment analysis, SVM

## I. INTRODUCTION

E-learning is an instructional system sent on the web, which gives understudies a chance to comprehend a given point completely. With the expansion being used of innovation in giving quality training, e-learning has been one of the strategies utilized by instructors to assist understudies with upgrading their abilities. E-learning helps in giving a uniform arrangement of training just as, in recording the learning example of the students. E-learning, to suit the necessities of various types of students is given as coordinated just as non-concurrent learning .Synchronous learning includes the course content being instructed with the assistance of live talks and video conferencing which enables the understudies to explain their questions quickly . In as opposed to this, offbeat learning comprises of pre-recorded recordings and practice meetings which the understudies can seek after at their own pace. To improve the nature of the course, a criticism and remarks segment is given in each course to comprehend the requirements of the understudies just as to see the openness of the course. These remarks go about as a helpful source to measure the nature of the courses, the convenience of the materials gave and the viability of the educator's instructing.

Assessment examination manages breaking down the human conduct by removing client feelings and sentiments from information which is as text. To get a knowledge into a huge arrangement of audits, we have to classify these surveys into nonpartisan, positive and negative remarks dependent on the substance of the surveys.

Backing Vector Machine is a managed AI calculation .It is a numerical model used to group the information into legitimate classes. One of the striking highlights of SVM is that it can learn regardless of the dimensionality of the component space. With the information plotted in the n dimensional space, the order of information is performed by finding the hyperplane that best separates the information focuses into various classes.

In this paper we utilize the criticism and audits gave in certain courses as a contribution to the supposition investigation calculation which separates the client's feelings and stores it as good, negative and impartial words. These arranged words go about as preparing information which is utilized by SVM to group the test information, which are the remarks from the remainder of the courses which we need to rate dependent on the input remarks into these three classes.

## II. LITERATURE REVIEW

Present day processing hardware is currently accessible to practically all degrees of instruction. The coming of Internet and innovation has to a great extent influenced the understudy network. To portray the viability of the innovation, Technology Acceptance Model is utilized as clarified by Edith et al.,2011 .Two primary develops, saw helpfulness and saw usability are utilized to depict this model. Seen Usefulness portrays the understudy's longing to utilize the innovation to upgrade their own exhibition and saw convenience is the straightforwardness that innovation gives to understudies, helping them to adapt freely [2]. Wolfgang et al., have done a survey on e-learning stages accommodated measurements. It has helped in finding the way that understudies of insights lean toward printed material, for example, books to electronic media, to improve their insight [1].E-learning sites give a criticism gathering and the surveys of understudies for a specific course and aides in checking the presentation of the instructor. Understudy conclusions can be named positive or negative by utilizing Hidden Markov Model or Support Vector Machine. Balaji et al.,2014 characterized the information by utilizing pre-handling of information, include extraction, utilization of SVM and HMM for precision and utilizing classifier mix rules for upgraded accuracy.[3]Sentiment Analysis encourages us comprehend a person's perspective communicated as text and consequently arrange it as certain and contrary. Zenia et.al., 2017 have shown the utilization of Naïve Bayes

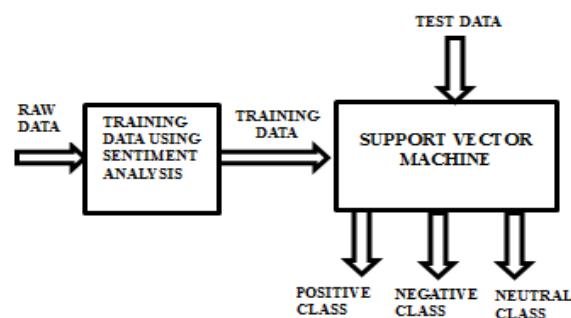
grouping calculation, Decision Tree and Support Vector Machine (SVM) for the characterization of adjusted audits of any client [5]

Interpersonal organizations permit the imparting of insights on a wide scope of issues by a large number of individuals. This can fill in as a rich information hotspot for notion examination and supposition mining .Sentiment investigation includes the investigation of assessments and feelings communicated in the literary structure. The primary goal being recognizing positive from the negative feelings passed on through remarks and appraisals. Francesco et al., have utilized Latent Dirichlet Allocation (LDA), a probabilistic methodology as a Sentiment grabber. This is utilized to build a chart dependent on the information area. The paper portrays these diagrams containing unmistakable weighted word sets. These word sets help in segregating words for slant characterization. This technique has been utilized on information from e-learning stages and causes the educators to tweak their showing dependent on the input of the understudies. [7]

### III. SUPPORT VECTOR MACHINE

SVM which is a directed AI calculation which can be utilized for grouping just as for relapse purposes however is typically utilized for the previous explanation. Every information thing is isolated into the classes which is finished by the hyperplanes. It tends to be utilized in high dimensional spaces successfully in any event, when the quantity of tests is lesser than the quantity of measurements. SVM is prepared by taking care of it archives which have been arranged previously. It functions admirably even on include sets which contain immense measure of information as a definitive goal is to just group the information, making it appropriate for text arrangement

#### CLASSIFICATION OF TEXT



**Figure 1: Classification of text as positive, Negative and Neutral**

Assumption examination is utilized to arrange text as certain, negative and unbiased dependent on the feelings passed on in the surveys. Contribution to the estimation investigation calculation

comprises of audits gathered from various courses in the e-learning sites, for example, Coursera and Udemy, which are given by European and Asian Universities. The estimation examination calculation takes the crude information for example audits and remarks and structures the preparation information (handled crude information) . Handling the crude information includes expulsion of stop words, accentuation checks and numbers. Extraordinary words from every one of these classifications are extricated and a preparation informational collection containing these exceptional words is framed.

The preparation information goes about as a learning worldview for the Support Vector Machine and causes it group the test information into positive, negative and impartial classes. The positive class is indicated by 0, negative by 1 and nonpartisan by 2. This document containing the arrangement done utilizing preparing informational index, alongside the test records is taken care of into SVM. Test documents contain surveys of courses gave by European and Asian Universities. The exceptional words contain positive, negative and unbiased words taken from test documents and are utilized to figure the level of each class for each test information record for examination reason. This correlation encourages us to investigate whether courses gave by Asian colleges are better or those gave by European Universities are better.

#### IV. ANALYSIS

For this study, reviews are taken from well-known E-learning websites Coursera and Udemy and fscopied into the text files 1.txt, 2.txt ,3.txt and 4.txt . A part of training data is given below in Table 1. Class 0, represents positive, class 1 represents negative and class 2 represents neutral words.

**Table 1. Training data extracted from reviews**

Unique words	Class
Okay	2
Difficult	1
good	0
Easy	0
Alright	2
Amazing	0
Awesome	0
Fine	2
Horrible	1
Okay	2
Terrible	1

Test files labelled 1.txt and 2.txt contain reviews collected from e-learning courses provided by Asian Universities and text files labelled 3.txt and 4.txt contain reviews collected from e-

learning courses provided by European Universities. A part of the Test files used in this study are as shown below.

**Figure 2: Test file 1.txt**

```

1.txt x 4.txt x 3.txt x 2.txt x
1 First time i had done any course online and this was a wonderful experience\n\nI learned all the stuff and the explanation provided was
  good would like to do more courses from here thank you!
2 Some more problems would have include for better hands-on with the launqege .
3 nice
4 Just satisfactory
5 Nothing to complain about
6 Extremely basic but Dr. Chuck is so interesting that you never get bored. You can pass the course in a few days with minimum or even no
  programming knowledge. It's a good starting point for beginners.
7 Great learning, from basic, I an a programmer, It was easy for me to understand concepts and syntax quickly .
8 Not very useful
9 It is an excellent course, a bit basic though
10 Quick win for whoever knows Java or C++.Try jumping onto the videos and complete assignments.
11 So much to finish after this course

```

**Figure 3: Test file 2.txt**

```

1.txt x 4.txt x 3.txt x 2.txt x
1 Course content may be good but the site performance is pathetic.Hangs and crashes many times. I have tried with IE and Chron. Also, on
  different laptops.
2 Nice Experience
3 Completely dissatisfied
4 NOT GOOD
5 It wasn't clear how the case study we were meant to look at related to what we had been, or were about to be, taught. I suspect I would
  have struggled with it and actually didn't bother.
6 I should say this is the worst experience in my life
7 Great course for beginners
8 Comprehensive and convenient.
9 It is one of the best courses to start with. The course content was very well presented and I loved learning from it.
10 Expected more things
11 This course is amazing for beginners , you will find almost everything

```

**Figure 4: Test file 3.txt**

```

1.txt x 4.txt x 3.txt x 2.txt x
1 good course
2 Great Course for beginners.
3 I was just revising my python skills while I came across this course, and have to say this is one of the best courses out there to learn.
  The professor goes slow for beginners and intermediates can easily go at 2x speed.
4 Great learning, from basic, I an a programmer, It was easy for me to understand concepts and syntax quickly .
5 Nice course
6 Good course for beginners. Has a few exercises, more could be added. The level of the course is very basic(just what the nene says).
  Overall it was a fun course, I learned a lot about Python and its syntax, looking forward to learning more pythoin in the days to come.
7 Amazing course
8 This is a very good course for basic python learners
9 thank u
10 I enjoyed the course. The concepts were explained in a simple and concise manner. The study material provided also helps if you want to
  study offline. DeFinitely shows the amount of hard work put in by the instructor. The chat section is also responsive and interactive. All
  my doubts were answered by awesome people in this community. Topics like DDP and generators, which can be really frustrating to go through
  as a beginner, were taught in the best way possible for a beginner.
11 This is a great course, the professor is passionate about Python which makes it easier to learn complex problems.
12 Awesome teacher

```

**Figure 5: Test file 4.txt**

```

1.txt      4.txt      1.txt      2.txt
1 Excellent Course! Easy to understand, clear explanations, informative, very well organized. I always wanted such type of training. The
  Instructor is very prompt in replying to doubts. The course is very well designed, practical and valuable for anyone who wants to enhance
  their skills. Thank you.
2 amazing learning experience
3 Great for a beginner learning a lot
4 Nice course for beginners
5 all the basics has been very well covered.
6 Expert Teaching and Training
7 Excellent course for beginners! Great job.
8 nice lecture and very helpful. thank you
9 Clear understanding of the topic covered and very helpful
10 good!
11 nice
12 Please share and discuss your views
  
```

The count of words belonging to each file and the total number of unique words in each file is stored, as shown below in Table 2

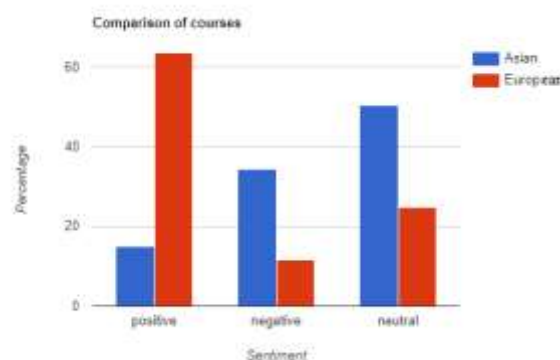
**Table 2. Test file classification as positive, negative and neutral**

File name	class	positive	negative	neutral	Total unique words
1.txt	2	204	198	734	1136
2.txt	1	106	509	300	915
3.txt	0	609	187	386	1182
4.txt	0	909	90	207	1206

Once the count of words belonging to each class is found, the percentage of positive, negative and neutral words is calculated for Europe and Asian Universities' courses. This calculation is shown in Table 3.

**Table 3: Comparison of courses**

University	Positive	Negative	Neutral
Asian	15.11	34.47	50.41
European	63.56	11.59	24.88



**Figure 6: Comparison between European and Asian Courses**

The above graph depicts the comparison between the Asian and European Universities' courses and from this comparison we can infer that courses provided by European Universities are preferred over Asian Universities' courses.

## V. CONCLUSION AND FUTURE WORK

Opinion examination and Support Vector Machine calculations have been utilized to play out an investigation on the surveys of courses gave by European and Asian Universities. In view of the above perceptions we can infer that European Universities' courses are favored over Asian Universities' courses. Hence helping understudies settle on an insightful choice with regards to picking on the web courses.

This investigation has great extension later on as it tends to be stretched out to dissect the variables adding to the determination of a specific course. This can assist colleges with improving their course content as per the client's need.

## VI. REFERENCES

1. Wolfgang Härdle, Sigbert Klinke, Uwe Ziegenhagen, "e-Learning Statistics -A Selective Review", Institute for Statistics and Econometrics, Humboldt-Universität zu Berlin, Germany.
2. Edith Galy, Clara Downey, Jennie Johnson, "The Effect of Using E-Learning Tools in Online and Campus-based Classrooms on Student Performance", Journal of Information Technology Education, Volume 10, 2011
3. Balaji Jagtap 1 and Virendrakumar Dhotre, "SVM and HMM Based Hybrid Approach of Sentiment Analysis for Teacher Feedback Assessment", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 3, Issue 3, May-June 2014, ISSN 2278-6856
4. Aakanksha, Er. Dinesh Kumar, "A HYBRID APPROACH FOR TEXT CLASSIFICATION USING HMM, SVM AND GENETIC ALGORITHM", International Journal For Technological Research In Engineering Volume 1, Issue 12, August-2014 ISSN (Online): 2347 – 4718
5. Zeenia Singla, Sukhchandan Randhawa, Sushma Jain, "Sentiment Analysis of Customer Product Reviews Using Machine Learning", 2017 International Conference on Intelligent Computing and Control (I2C2)
6. Andrew L. Maas, Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, Christopher Potts, "Learning word vectors for sentiment analysis", Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics, pages 142–150, Portland, Oregon, June 19-24, 2011. © 2011 Association for Computational Linguistics
7. Francesco Colace, Massimo De Santo, Luca Greco, "SAFE: A Sentiment Analysis Framework for E-Learning", Università degli Studi di Salerno, Fisciano (Salerno) – Italy
8. Fabio Clarizia, Francesco Colace, Massimo De Santo, Marco Lombardi, Francesco Pascale, Antonio Pietrosanto, "E-Learning and Sentiment Analysis: A Case Study" DIIn, University of Salerno via Giovanni Paolo II, 132, Fisciano (SA), Italy
9. Muhammad Zubair Asghar, Shakeel Ahmad, Afsana Marwat, Fazal Masud Kundi, "Sentiment Analysis on YouTube: A Brief Survey", Institute of Computing and Information Technology (ICIT), Gomal University, D. I. Khan, Pakistan, Faculty of Computing and Information Technology in Rabigh (FCITR), King Abdul Aziz University (KAU) Saudi Arabia.



10. Zied Kechaou, Mahmoud bouhajja ben ammar, Adel M. Alimi, “ Improving e-learning with sentiment analysis of users' opinions”, 2011 IEEE Global Engineering Education Conference (EDUCON) – "Learning Environments and Ecosystems in Engineering Education"
11. Badreddine Benyacoub, Souad ElBernoussi, Abdelhak Zoglat, EL Moudden Ismail, “Classification with Hidden Markov Model”, Applied Mathematical Sciences, Vol. 8, 2014, no. 50, 2483 – 2496
12. Yasemin Altun, Ioannis Tsochantaridis, Thomas Hofmann, “ Hidden Markov Support Vector Machines”, Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003), Washington DC, 2003.
13. Ahmad, A.R, Viard-Gaudin, C., Khalid, M, “Lexicon-based Word Recognition Using Support Vector Machine and Hidden Markov Model”, 2009 10th International Conference on Document Analysis and Recognition
14. Shanthanu Chakrabartty, Gunnet singh and Gert Cauwenberghs, “Hybrid Support Machine / Hidden Markov Model approach for continuous speech recognition”, John Hopkins University, USA
15. Xiaoyi Zhao and Yukio Ohsawa , “ Sentiment Analysis on the Online Reviews Based on Hidden Markov Model”, *Journal of Advances in Information Technology Vol. 9, No. 2, May 2018*