EFFECTIVE MULTILAYERED MODEL FOR ANALYSIS OF STUDENTS BEHAVIOR USING SOCIAL MEDIA MINING

Mathews Emmanuel

Assistant Professor

Department of computer Applications,

Saint Josephs college of Engineering and Technology, Palai, India

hellomathews@gmail.com

Nimmy Chacko

Assistant Professor

Department of Basic Science

Amaljyothi college of Engineering and Technology, Koovappally, India

nimmichan87@gmail.com

ABSTRACT

Educational Data Mining or EDM refers to the unique and effective approach related to data mining, statistics and machine learning so that the predictions and analytics can be done in the stream of education and academic domain. A number of algorithms and research approaches are

implemented so far for this segment still there is scope of lots of research. In this research paper, the sentiment data analysis from twitter and other social media is projected and underlined in the domain of educational data mining. Using this approach, the students' behavior and their discussions on twitter, facebook, instagram and other related platforms can be evaluated using advance natural language processing algorithms. Using this methodology, the interests, patterns and diversions of students can be evaluated and then predicted about their interest in the classrooms and further career orientations.

Keywords - Educational Data Mining, Sentiment Data Analysis, Twitter Analytics, Social Media Analytics for Educational Data Mining

PREAMBLE AND INTRODUCTION

Educational Data Mining refers and associated with the tools and techniques for the automatically extraction of aspects from huge repositories of data sets generated or related with the user's activities on learning in the educational environment. This data is very extensive, precise and fine-grained for research and predictive analysis.
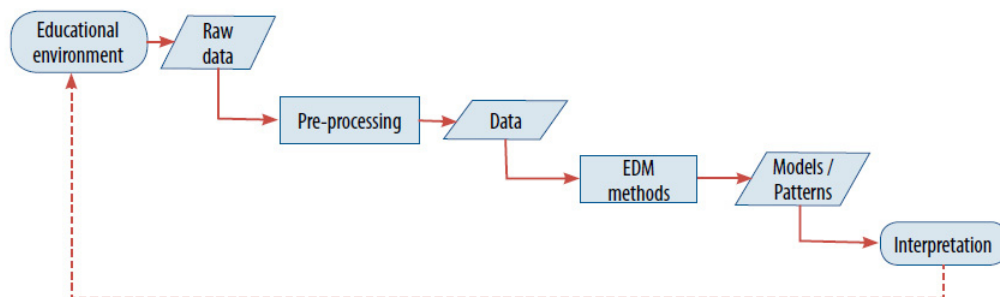

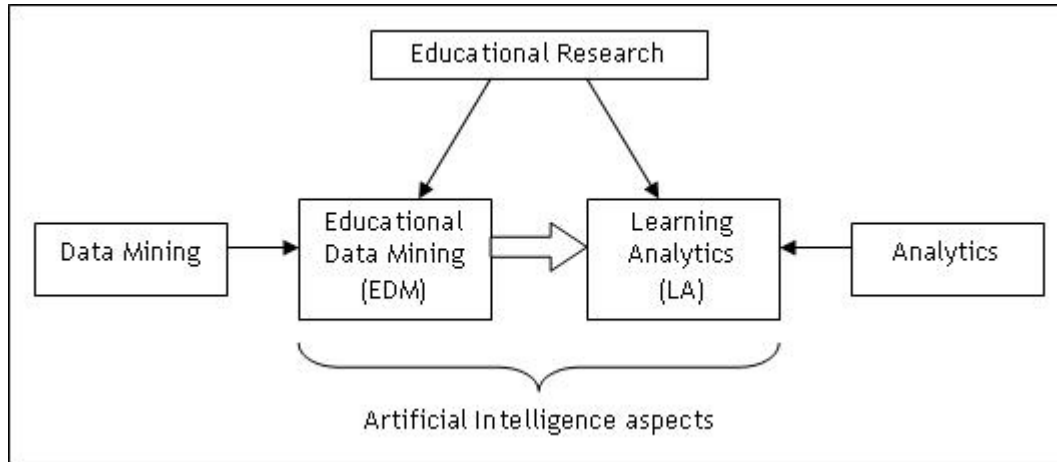Figure 1 – Education Data Mining and Related Aspects

Figure 2 – Artificial Intelligence Aspects with Education Data Mining

GOALS AND OBJECTIVES OF EDUCATIONAL DATA MINING

Yacef and Baker presented the following goals associated with EDM

1. Prediction of Students' Behavior in relation to future learning

2. Analysis of Models associated with domains

3. Study and Deep Investigation of effects of Educational Segment Support

4. Advancements in the Scientific Knowledge on Academia

USERS, STAKEHOLDERS AND ACTORS IN EDM

- Students / Learners

- Teachers / Mentors / Educators

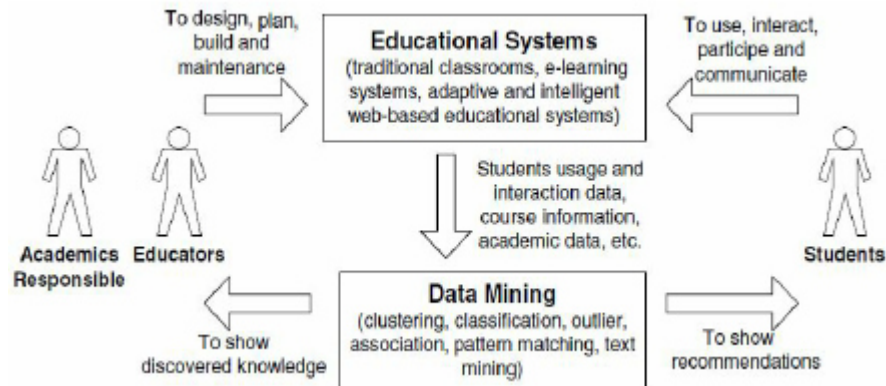- Researchers / Professionals

- Administrating Authorities

Figure 3 – Educational Systems and Academic Dimensions

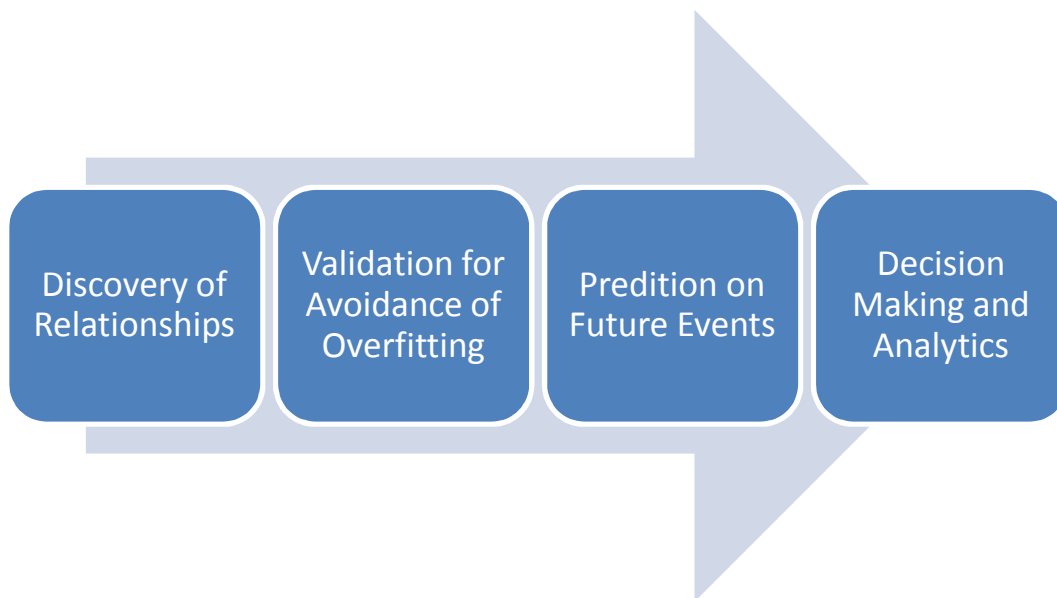## PHASES AND LAYERS IN EDUCATIONAL DATA MINING

Figure 4 – Phases and Layers of Educational Data Mining

APPLICATIONS OF EDM

- Visualization and Analysis of students' data

- Deep feedback and analytics

- Higher level Recommendations and predictions for students

- Prediction of students' performance and patterns

- Modeling of students

- Detection of undesirable and specific student behaviors

- Features Points Extraction from the Students' Datasets

- Grouping and Segmentation of students

- Social network analytics and Future Predictions

- Development of concept based maps for interrelationships

- Construction of courseware based on the student's interests and behavior

- Planning, scheduling and effective decision making

LIMITATIONS AND CHALLENGES OF EDM

- Generalization and Adaptability

- Privacy and Self Controlled Work

- Plagiarism and Copyright Issues

- Adoption and Global Acceptance

CLASSICAL METHODS FOR ANALYTICS

- Deep analysis of Contents

- Discourse Based Analytics
- Social Learning Based Analytics
- Disposition Based Analytics

PROPOSED APPROACH AND METHODOLOGY

In every university and educational organization, the live monitoring of tweets and messages can be done on the accounts of students and teachers so that their interests can be analyzed and predicted.

In addition, the separate and institution based social media platforms can be developed. Using this approach, the live discussions, messages and tweets can be investigated based on their interests, feature points and career orientations.

In the proposed algorithmic approach having multiple layers, the following components can be integrated for effective prediction of students' behavior in the classrooms and academic practices

- Fetching of Live Data Sets from Social Media using Advance Programming Languages
- Parsing of Data Sets to the understandable and process formats
- Implementation and Integration of Natural Language Toolkits for Sentiment Analysis
- Positive and Negative Tweets Investigation for analyzing the behavior and patterns of students' interests.
- Plotting of graphs and charts for effective predictions

Figure 5 – Creation of Developer Account for Educational Data Mining from Twitter

Figure 6 – Setting Up the Credentials and Keys for Educational Data Mining from Twitter
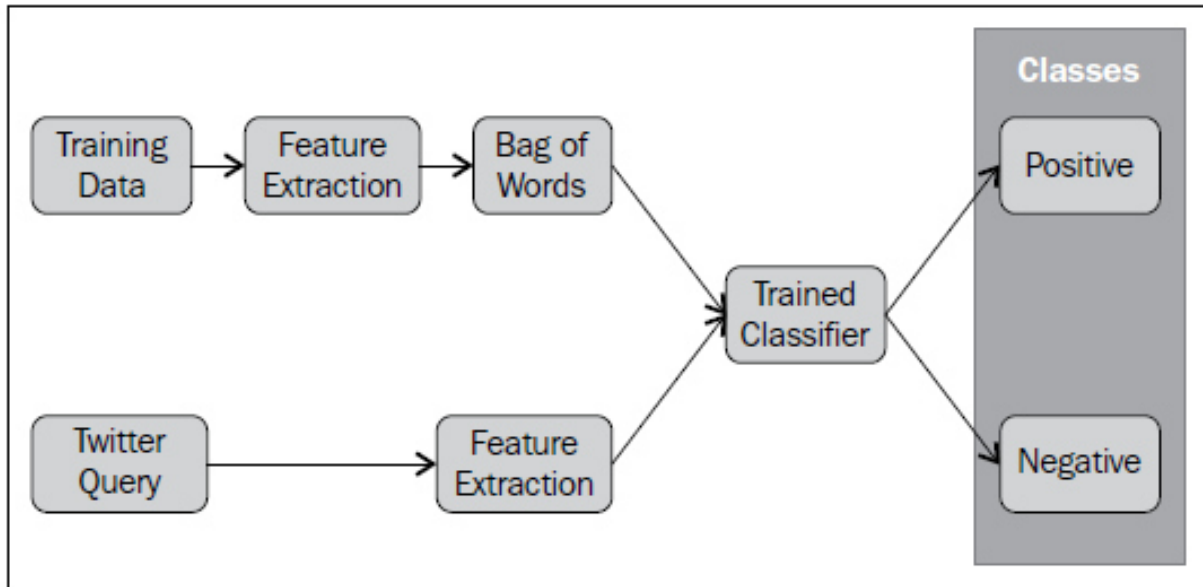
Figure 7 – Sentiment Dimensions and Flow for Educational Data Mining

- Fetching of Live and Real Tweets from Twitter using Twitter4J Advance Java Interface
- Fetching of genuine dataset in Engine Readable Format
- Parsing of the File Format to readale
- Extraction of Interesting / Feature Points from Tweets
- Implementation of the Tokenizing on Tweets
- Implementation of the Words / Token based the deep investigation algorithm
- Integration of Classical approach of sentiment analysis using frequency count.
- Comparison of results with the NLTK (Natural Language Toolkit) using Advance Java for the Efficiency and Performance Analysis
- Analytics using the Sentiment Score

- Plotting of the Graphs and Partitioning using the APIs of JFreeCharts
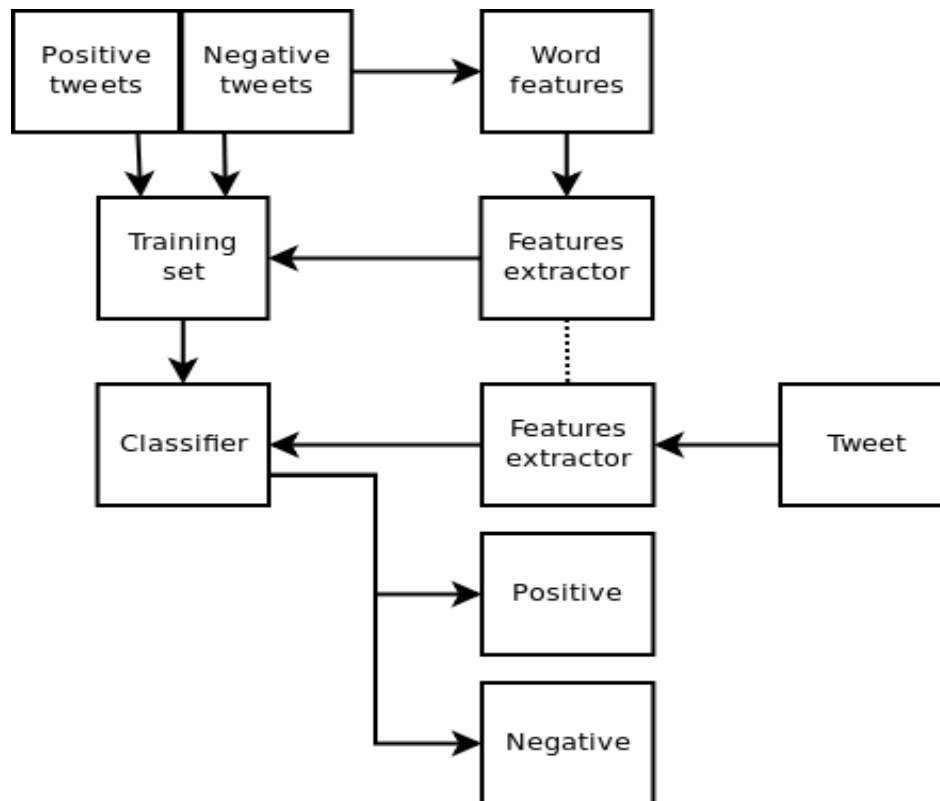
Figure 8 – Features Extraction and Classifiers in EDM

TOOLS AND TECHNOLOGIES USED EFFECTIVE PREDICTIONS AND ANALYTICS

- Advance Java
- Eclipse IDE
- Twitter4J

- MySQL Database Engine

- Tomcat Server

- Notepad++

- Dia - The Diagram Editor

- Stanford NLP API

- JFreeCharts JAVA APIs for Plotting of Graphs

SIMULATION RESULTS

In the following results, the behavior of students on the keywords "ALGORITHM DESIGN" is analyzed and predicted in real time after analytics of tweets of twitter.

From the results, it is found that this keyword is very prominent and discussed by the students as compared to the keyword "DISCRETE MATHEMATICS"
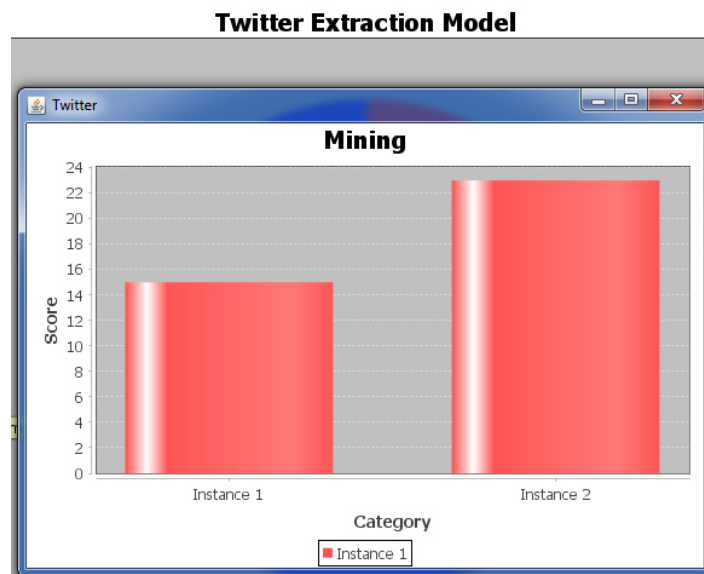
Figure 9 – Results from Twitter Analytics
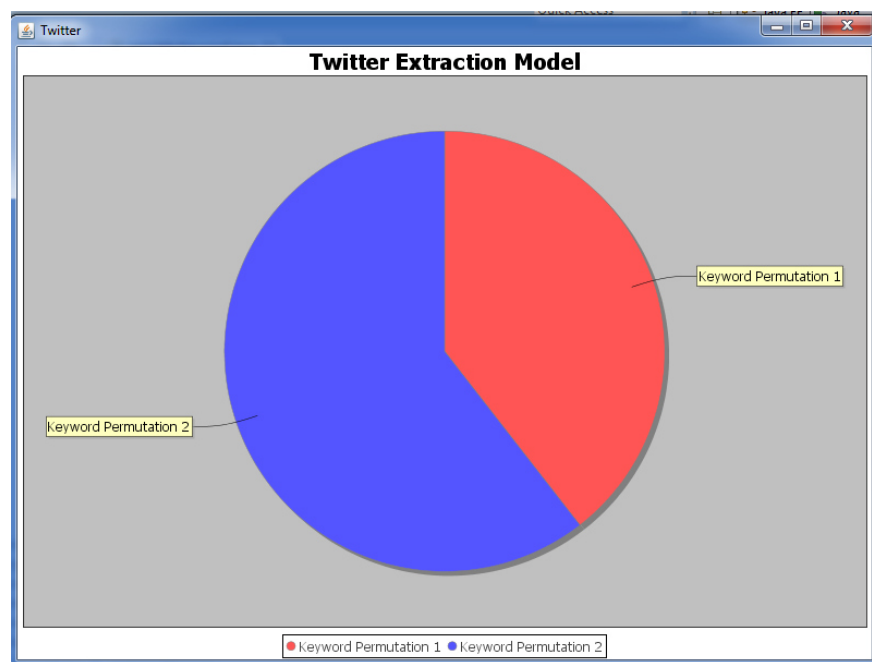
Figure 10 – Results from Twitter Analytics

Keyword Under Search -> algorithm

Keyword Under Search -> design

Keyword Size -> 1

Keyword Size -> 1

<-->

First Permutation -> algorithm design

Second Permutation -> design algorithm

<--->

Keyword Under Twitter Live Scan -> algorithm design

SAMPLE EXTRACTED LIVE TWEETS OF TWITTER ON STUDENT'S DISCUSSIONS

algorithm design #Marketing https://t…: null: null: UserJSONImpl{id=3227599383, name='Yvonne Tarasoua', screenName='dunsites',

22 14:48:07 IST 2015, favouritesCount=318,

6029248356692 82816/19oPvl_Q

withheldInCountries=null}: false: null [<Sentiment Score -> 1>] 711026308806004737 - Sat Mar 19 08:37:36 IST 2016: Design Diary no. 84 https://t.co/bqsZEMiTDw: null: null: UserJSONImpl{id=4173421091, name='Design Trends', screenName='Todays_Design', 664887391279120388/99WGx8zk_norm al.jpg',

withheldInCountries=null}: false: null [<Sentiment Score -> 1>] 711024837464985601 - Sat Mar 19 08:31:46 IST 2016: RT @Red_Web_Design: Working on Your #SEO? 5 Google Algorithm Updates You Need to Know About: https://t.co/DoUEjY0mYM #Marketing https://t…: null: null: UserJSONImpl{id=3227649389, name='Deann Moustaid', screenName='DeannMoustaid',

withheldInCountries=null}: false: null [<Sentiment Score -> 1>] 711023473326313473 - Sat Mar 19 08:26:20 IST 2016: RT @Red_Web_Design: Working on Your #SEO? 5 Google Algorithm Updates You Need to Know About: https://t.co/DoUEjY0mYM #Marketing https://t…: null: null: UserJSONImpl{id=3227622939, name='Glenna Brannon', screenName='BrannonGlenna',

null: null: UserJSONImpl{id=572820276, name='AlgorithmProblemBot', screenName='AlgorithmProble', withheldInCountries=null}: false: null [<Sentiment Score -> 1>] 711016232359747584 - Sat Mar 19 07:57:34 IST 2016: RT @Red_Web_Design: Google Hummingbird: An Update That Could Dramatically Affect Your SEO: https://t.co/NGgGaa2HAG #SEO https://t.co/ztPT…: null: null: UserJSONImpl{id=3426786869, name='Ben Stevens', screenName='BenStevo22', location='United Kingdom', description='Ex Military - Social media manager and Web Designer for https://t.co/RKZDDzcHk6 Get in touch to buy your domain in our shop and get your business started!', isContributorsEnabled=false, profileImageUrl withheldInCountries=null}: false: null [<Sentiment Score -> 1>] 711009175028031488 - Sat Mar 19 07:29:31 IST 2016: RT @AIGAdesign: Got 60 seconds? Use 'em to get up to speed on the top design news of the week: https://t.co/IXgNZBjbZ7 https://t.co/pRVfqNj…: null: null: UserJSONImpl{id=429417472, name='AIGA DC SHINE', screenName='AIGADC_SHINE', location='Washington DC', description='SHINE, An AIGA DC Mentoring Initiative', withheldInCountries=null}: false: null [<Sentiment Score -> 1>] 711004216718786561 - Sat Mar 19 07:09:49 IST 2016: RT @a2c_iwasaki: ???https://t.co/cPLkXmFCux ?https://t.co/hhDIlmz5kt??????????…: null: null: UserJSONImpl{id=210808882, name='Umepon', screenName='shunji_umetani', location='Osaka, Japan', description='Researcher, Modeling and algorithm in mathematical optimization, Heuristic algorithm for large-scale combinatorial optimization and integer program', withheldInCountries=null}: false: null [<Sentiment Score -> 1>] 711000007600697344 - Sat Mar 19 06:53:06 IST 2016: Golf Club Fitting Hardware / Software / Algorithm Development - Software Design Freelance Job #project https://t.co/k9hL5EjZut: null: null:

UserJSONImpl{id=2541593905, name='Freelance DP', screenName='freelance_dp', location='', description='freelance, contract work, 1099, freelancing, freelancers, contract employees', 475770936105246721/A7VGkZmK_nor mal.jpeg', withheldInCountries=null}: false: null [<Sentiment Score -> 1>] 710994348486238208 - Sat Mar 19 06:30:36 IST 2016: ???https://t.co/cPLkXmFCux?https://t.co/hhDIlmz5kt??????????????????????????: null: null: UserJSONImpl{id=284008616, name='Atsushi Iwasaki', screenName='a2c_iwasaki', location='????????st', description='Associate Professor, University of Electro-Communications, Hanshin-fan. Informatics and Economics: optimization and game theory', withheldInCountries=null}: false: null [<Sentiment Score -> 1>] 710993829659328513 - Sat Mar 19 06:28:33 IST 2016: Small Business Mobile Website Design Alert! Google announces new algorithm update for non mobile friendly websites. https://t.co/xnQw5Xwyjm: null: null: UserJSONImpl{id=1625770524, name='Buy Cars Right', screenName='buycarsright', location='USA', description='Get important tips on buying new and used cars and save time and money. Find New and Used Cars For Sale.', withheldInCountries=null}: false: null [<Sentiment Score -> 1>] 710993828984066048 - Sat Mar 19 06:28:33 IST 2016: Small Business Mobile Website Design Alert! Google announces new algorithm update for non mobile friendly websites. https://t.co/8etyljAQOu: null: null: UserJSONImpl{id=404880801, name='Autonet Media', screenName='AutonetMedia', location='Phoenix Arizona', description='Automotive Content Marketing company providing info to Car Buyers and sellers. Our network of websites offers car buying help, and new and used car classifieds.',

CONCLUSION AND SCOPE OF FUTURE WORK

Using Sentiment Analysis Integration, the effective results can be taken for educational analytics. Still there is huge scope of research. Following points can be integrated for higher accuracy and integrity in the results.

- Identification of Videos and Special Files
- Advance Prediction Techniques
- Plotting and Behavior Analysis
- Deep Extraction of Sentiments
- Association Mining for getting relationship between the tweets
- Integration of Advance Data Mining Approaches

REFERENCES

[1] Scheuer, O., & McLaren, B. M. (2012). Educational data mining. In *Encyclopedia of the Sciences of Learning* (pp. 1075-1079). Springer US.

[2] Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *JEDM-Journal of Educational Data Mining*, *1*(1), 3-17.

[3] Siemens, G., & d Baker, R. S. (2012, April). Learning analytics and educational data mining: towards communication and collaboration. In*Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252-254). ACM.

[4] Wen, M., Yang, D., & Rosé, C. P. (2014). Sentiment analysis in MOOC discussion forums: What does it tell us. *Proceedings of educational data mining*, *1*.

[5] Bifet, A., & Frank, E. (2010, October). Sentiment knowledge discovery in twitter streaming data. In *Discovery Science* (pp. 1-15). Springer Berlin Heidelberg.

[6] Romero, C., Ventura, S., Pechenizkiy, M., & Baker, R. S. (Eds.). (2010).*Handbook of educational data mining*. CRC Press.

[7] Baradwaj, B. K., & Pal, S. (2012). Mining educational data to analyze students' performance. *arXiv preprint arXiv:1201.3417*.