Naive Bayes Model for Labeling an Approach in the Trade Center

M. Suhela Mariyam[1], K. Geetha. M. E[2]

Second Year Mca[1], Assistant Proffessor[2]
Department Of Computer Science And Applications
Faculty Of Computer Science And Engineering
Periyar Maniammai Institute Of Science And Technology
Vallam, Thanjavur, Tamil Nadu, India.

ABSTRACT:
The Trading Center hasthe volume of the approach is more, compared to previous.So it is difficult to analyse the approach of the entity by manual process and also faces a lot of complications while analysing the approach, for classifying the document of approach. Initially there is no label for the approach to launch for the visit, but label plays an important role in Trade Center while launching the product.In previous, analysing the unstructured data for labelling is done.But it has some drawbacks while using the algorithm. For example,while using joint and alignment algorithm we cannot simultaneously use the consensus and complementary properties. To overcome, In this project we proposed some algorithm for labelling the approach for the product to launch. When the entity has uploaded the document of approach in the Trade Center, first of all it just analysed by the super admin and extract the label for the respective document. The analysed label is used for categorising the approach and also for scheduling it. When the label has found then admin can check for the profit or loss while launching the product.

I.   INTRODUCTION:
In Trade Center for analyzing the approach for the product by the document classification is by using an specific algorithm. By using the algorithm admin can find the label for the document of approach. By the finding the label, further process has been taken to launch the approach. Here result of the approach is also analyzed, which is in the image format. Admin is just analyzing the document for launching the approach in Trade Center. This launched approach is seen by the registered client in the Trade Center. But before launching the product, the prediction for the live launching product for analyzing that the product can be a profit or loss in percentage after the product has launched. After labelling, admin will predict for the live launching product to be profit or loss after the product has launched. The Prediction for the profit or loss percentage is analyzed by based on the given conditions in the algorithm. After the prediction process the scheduling for the approach has takes place. Once the scheduling process has completed the percentage details will be sent to the respective entity.

EXISTING SYSTEM:
In the classification, previous system uses the joint and alignment representation learning method that cannot uses the consensus and complementary properties of multi view data for learning the shared and specific representations.
To make use of the consensus and complementary properties as well as shared and specific representations, have to use the shared and specific representation learning network.

DISADVANTAGES OF EXISTING SYSTEM:
✓ Not cost-efficient
✓ Output of label can be change.
✓ More algorithms are used for labeling.
✓ To increase accuracy more recursions are used
✓ Can't be used in wide applications.

PROPOSED SYSTEM:
The Trade Center project has the document classification which is for labelling the document of approach. This document classification is done by the specific algorithm which is to label the document of approach. The specific algorithm used for the document

classification is Naïve Bayes, which classifies the document based on the content in the document of approach.

After labelling, admin will predict for the live launching product to be profit or loss after the product has launched. The Prediction for the profit or loss percentage is analyzed by based on the given conditions in the algorithm.After the prediction process the scheduling for the approach has takes place. Once the scheduling process has completed the percentage details will be sent to the respective entity.

ADVANTAGES OF PROPOSED SYSTEM:
✓ It has the responsive interaction between the entity and the admin
✓ Its improves time efficiency by using classification rather than the manual
✓ It increases the accuracy for labelling the approach.
✓ It reduces the man power by using the algorithm
✓ It neglects the human error while analyzing the document

II. SYSTEM DESIGN:
2.1 Modules:
1) Entity
2) Client
3) Admin
4) Super Admin

2.2 Module Description:
1.Entity:
This module first gives the login page to go for the entity home page and also it gives the link for the registration page for the entity to register the details where they can give the name, email id, contact number, password and confirm password in this, password and confirm password should be same. After login process entity get the entity home page and in that there are three sub modules such as upload, intimations and logout. First module is upload where the upload process will be takes place. In the upload process entity is just uploading the model type, space or required seats, Date to be held, document file which has the approach of entity and image file which has the model of the approach. Once entity fills and just uploaded this data will goes to super admin. Second module is intimations where it consists of the details about the title for the approach, computer generated code for the approach, model type that permitted for that product, space or required seats for the meeting and shows the date and time scheduled for the meeting and also

displays the predicted percentage value for the approach to be profit or loss.
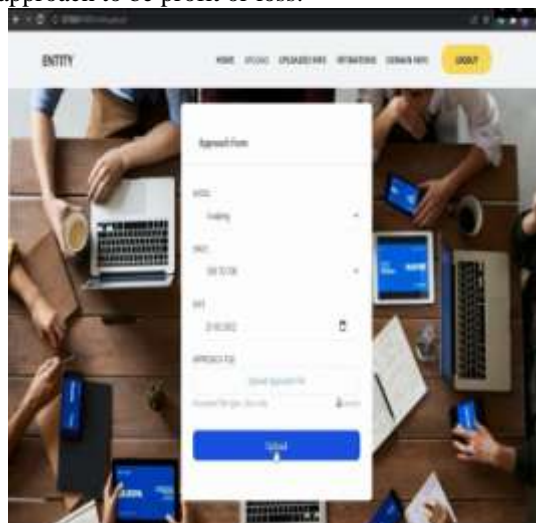
2.Client:
In this module initially it gives the login page with the registration page link where in that login for the client process takes place and in registration page the client will give the details to register for the trading process. The details while the client is registering are name of the client, email id, contact number, password and confirm password. In this, password and confirm password should be same while registering. After the login process the page will redirect to the client home page which has the four modules such as register, ticket, visit and logout. In first module of register has the register form for the client to register for the meeting which is launched by super admin and in that form has product code, model type, date and time which is launched and also shown in home page. Once client has registered for this visit, the details has sent to the admin for the approval for the visit and also for ticket number. Second module named ticket consist of the table which shows the label, product code for which the client registered, model type, date and time and also shows the ticket number for the visit. Third module named visit consists of the form where the client can enter into the meeting which already the client have registered for. Fourth sub module in a client is logoutwhere the client can logout from their logged in page where the session details are erased.

3.Admin:

This module gives out the login page initially for the admin to login into their particular account. After login process the page is redirected to the admin home page where it gives four sub modules such as approve, ticket, schedule and logout. First module named approve is used to show the table consists of the register queue of client where the client is registered that data will be stored in this table. From that table admin can verify the client registration details and can approve for the further process for the client. Second sub module named ticket where the table in this module will show the registered details of client for the visit or meeting. In this admin can approve the registered details for a visit and also ticket has been generated randomly and also send to the client, that means the particular client can participate in a visit with the ticket generated.Third sub module named scheduled where theanalyzedapproach entity details will be shown and also admin can schedule for launching purpose. In this table the schedule button is active when the analyzed approach has predicted by super admin otherwise it will be in in active mode. Once the admin uses the schedule button then the page is redirected to the scheduling form where that particular entity details have filled in the input field where admin can change according to the time management. Once the details are verified in scheduling form by clicking the schedule button admin can schedule the approach to expose in a visit. Fourth sub module named logout is used for the admin to logout from the admin page.

4.Super Admin:

In this module initially it gives the login page the super admin may login to super admin page. After a successful login it will redirected to the super admins home page. This page has the five sub modules such as uploaded, predict, launch, display and logout. First sub module named uploaded which gives out the table with the name of the entity, model type, date, space and also with approach document file and image file. Here super admin can analyze the file for labeling the approach. By using the analyze button super admin can start analyze. Second sub module named predict which shows the table consists of the data which is to be predicted. Once analyzing process completed then the resultant data will be transferred to this table. While using the predict button,prediction process for the labeled approachthat results in the percentage of profit or loss for the approach by comparing with the last year launched approach. When the prediction process completed then the schedule button in the sub module of schedule present in the admin module will be active and also gives the predicted details in the predicted field. Third sub module named launch will gives the details of scheduled approach that means the details in the tables are ready to launch in the home page. When super admin uses the launch button, then the details in the records are launched and that is displayed in the display table. If super admin wants to delete the post which is launched, then they can use the delpost button. Fourth sub module named display which displays the live launching details. If super admin delete the post then that particular post will be deleted in this table. Fifth sub module named logout which is used for the super admin to logout off from the super admin page.

III. SYSTEM TESTING:

3.1 Software testing:
Software testing is a critical element of software quality assurance and represents the ultimate review of specification, design and coding. In fact, testing is the one step in the software engineering process that could be viewed as destructive rather than constructive.

3.2 Unit Testing:
Unit testing focuses verification effort on the smallest unit of software design, the module. The unit testing we have is white box oriented and some modules the steps are conducted in parallel.

1. WHITE BOX TESTING:
This type of testing ensures that
➤ All independent paths have been exercised at least once
➤ All logical decisions have been exercised on their true and false sides
➤ All loops are executed at their boundaries and within their operational bounds
➤ All internal data structures have been exercised to assure their validity.

To follow the concept of white box testing we have tested each form .We have created independently to verify that Data flow is correct, All conditions are exercised to check their validity, All loops are executed on their boundaries.

2. BASIC PATH TESTING:
The established technique of flow graph with Cyclamate complexity was used to derive test cases for all the functions. The main steps in deriving test cases were:
Use the design of the code and draw correspondent flow graphs.
Determine the Cyclamate complexity of the resultant flow graph, using formula:
$V (G) = E-N+2$ or

$V (G) = P+1$ or
$V (G) =$ Number of Regions
Where V (G) is Cyclomatic complexity,
E is the number of edges,
N is the number of flow graph nodes,
P is the number of predicate nodes.
Determine the basis of set of linearly independent paths.

3. CONDITIONAL TESTING
In this part of the testing each of the conditions were tested to both true and false aspects. And all the resulting paths were tested. So that each path that may be generated on particular condition is traced to uncover any possible errors.

4. DATA FLOW TESTING
This type of testing selects the path of the program, according to the location of the definition and use of variables. This kind of testing was used only when some local variable were declared. The definition-use chain method was used in this type of testing. These were particularly useful in nested statements.

5. LOOP TESTING
In this type of testing all the loops are tested to all the limits possible. The following exercise was adopted for all loops:
➤ All the loops were tested at their limits, just above them and just below them.
➤ All the loops were skipped at least once.
➤ For nested loop test the innermost loop first and then work outwards.
➤ For concatenated loops the values of dependent loops were set with the help of a connected loop.

IV. SYSTEM SECURITY:
Security system can be divided into four related issues: The protection of computer based resources that includes hardware, software, data, procedures and people against unauthorized use or natural.
Disaster is known as System Security.
➤ Security
➤ Integrity
➤ Privacy
➤ Confidentiality

SYSTEM SECURITY refers to the technical innovations and procedures applied to the hardware and operation systems to protect against deliberate or accidental damage from a defined threat.
DATA SECURITY is the protection of data from loss, disclosure, modification and destruction.
SYSTEM INTEGRITY refers to the power functioning of hardware and programs, appropriate

physical security and safety against external threats such as eavesdropping and wiretapping.

PRIVACY defines the rights of the user or organizations to determine what information they are willing to share with or accept from others and how the organization can be protected against unwelcome, unfair or excessive dissemination of information about it.

CONFIDENTIALITY is a special status given to sensitive information in a database to minimize the possible invasion of privacy. It is an attribute of information that characterizes its need for protection.

4.1 SECURITY IN SOFTWARE:
System security refers to various validations on data in the form of checks and controls to avoid the system from failing. It is always important to ensure that only valid data is entered and only valid operations are performed on the system. The system employs two types of checks and controls:

1.CLIENT SIDE VALIDATION
Various client side validations are used to ensure on the client side that only valid data is entered. Client side validation saves server time and load to handle invalid data. Some checks are imposed:

➤ JavaScript in used to ensure those required fields are filled with suitable data only. Maximum lengths of the fields of the forms are appropriately defined.

➤ Forms cannot be submitted without filling up the mandatory data so that manual mistakes of submitting empty fields that are mandatory can be sorted out at the client side to save the server time and load.

➤ Tab-indexes are set according to the need and taking into account the ease of use while working with the system.

2.SERVER SIDE VALIDATION
Some checks cannot be applied on the client side. Server side checks are necessary to save the system from failing and intimating the user that some invalid operation has been performed or the performed operation is restricted. Some of the server side checks imposed is:

➤ A server side constraint has been imposed to check for the validity of primary key and foreign key. A primary key value cannot be duplicated. Any attempt to duplicate the primary value results in a message intimating the user about those values through the forms using foreign key can be updated only of the existing foreign key values.

➤ The user is intimated through appropriate messages about the successful operations or exceptions occurring at server side.

➤ Various Access Control Mechanisms have been built so that one user may not agitate upon another. Access permissions to various types of users are controlled according to the organizational structure. Only permitted users can log on to the system and can have access according to their category. User- name, passwords and permissions are controlled the server side.

➤ Using server side validation, constraints on several restricted operations are imposed.

V. CONCLUSION:
In this project Labelling for the document of approach has a main role. So here labellingfor a document is done by using the Naïve Bayes. In this the approach can be submitted by the document. But in future the approach is directly fed into web so admin can perform the web classification for the approach to be launched. The purpose for upgrading future enhancement to web classification is much more effiecient than the document

VI. FUTURE WORK:
In the future, Because in document classification there are more number of format in the document but have to specify the extention to verify. But in web classification there is no need for the extension or some other drawbacks. The entity can directly give the approach in the web application where there has no need of format and can classify the approach through web classifications.

REFERENCE:
[1] J. Zhao, X. Xie, X. Xu, and S. Sun, "Multi-view learning overview: Recent progress and new challenges," Information Fusion, vol. 38, pp. 43–54,2017.
[2] C.Xu,D. Tao,andC.Xu,"Asurveyonmulti-viewlearning,"arXivpreprint arXiv:1304.5634,2013.
[3] Y. Li, M. Yang, and Z. M. Zhang, "A survey of multi-view representation learning," IEEE Transactions on Knowledge and Data Engineering, vol. 31, no. 10, pp. 1863–1883,2019.
[4] T.Baltrušaitis,C.Ahuja,andL.-P.Morency,"Multimodalmachine learning: A survey and taxonomy," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 41, no. 2, pp. 423–443,2019.
[5] X. Xie and S. Sun, "Multi-view support vector machines with the consensus and complementarity information," IEEE

Transactions on Knowledge and Data Engineering,2019.

[6] C. M. Christoudias, R. Urtasun, and T. Darrell, "Unsupervised feature selection via distributed coding for multi-view object recognition," in CVPR, 2008, pp.1–8.

[7] T. Zhou, C. Zhang, C. Gong, H. Bhaskar, and J. Yang, "Multiview latent space learning with feature redundancy minimization," IEEE Transactions on Cybernetics, pp. 1–14,2018.

[8] H. Tao, C. Hou, F. Nie, J.Zhu, and D. Yi, "Scalable multi- view semi-supervised classification via adaptive regression,"IEEE Transactions on Image Processing, vol. 26, no. 9, pp. 4283–4296,2017.

[9] N.SrivastavaandR.R.Salakhutdinov,"Multim odallearningwith deep boltzmann machines," in NeurIPS, 2012, pp.2222–2230.

[10] F. Nie, G. Cai, J. Li, and X. Li, "Auto-weighted multi-viewlearning for image clustering and semi-supervised classification," IEEE Transactions on Image Processing, vol. 27, no. 3, pp. 1501–1511,2018.

[11] Y. Cheng, X. Zhao, R. Cai, Z. Li, K. Huang, and Y. Rui, "Semi- supervised multimodal deep learning for rgb-d object recogni- tion." in IJCAI, 2016, pp.3345–3351.

[12] L. Zhao, Q. Hu, and Y. Zhou, "Heterogeneous features integration via semi-supervised multi-modal deep networks," in ICONIP, 2015, pp.11–19.

[13] J. Li, J. Zhao, F. Zhao, H. Liu, J. Li, S. Shen, J. Feng, and T. Sim, "Robust face recognition with deep multi-view representation learning," in ACM MM, 2016, pp.1068–1072.

[14] Z. Zuo, Y. Luo, D. Tao, and C. Xu, "Multi-view multi-task feature extraction for web image classification," in ACM MM, 2014, pp. 1137–1140.

[15] X. Chen, S. Chen, H. Xue, and X. Zhou, "A unified dimensionality reduction framework for semi-paired and semi-supervised multi-view data," Pattern Recognition, vol. 45, no. 5, pp. 2005–2018,2012.

[16] X.-Y. Jing, F. Wu, X. Dong, S. Shan, and S. Chen, "Semi-supervised multi-view correlation feature learning with application to web- page classification," in AAAI, 2017, pp.1374–1381.

[17] G. Andrew, R. Arora, J. Bilmes, and K. Livescu, "Deep canonical correlation analysis," in ICML, 2013, pp.1247–1255.

[18] W.Wang,R.Arora,K.Livescu,andJ.Bilmes,"O ndeepmulti-view representation learning," in ICML, 2015, pp.1083–1092.

[19] V. Noroozi, S. Bahaadini, L. Zheng, S. Xie, W.Shao, andP. S. Yu, "Semi-supervised deep representation learning for multi-view problems," in Big Data, 2018, pp.56–64.

[20] A. Frome, G. S. Corrado, J. Shlens, S. Bengio, J. Dean, T. Mikolovet al., "Devise: A deep visual-semantic embedding model," in NeurIPS, 2013, pp.2121–2129.

[21] D. Harwath, A. Torralba, and J. Glass, "Unsupervised learning of spoken language with visual context," in NeurIPS, 2016, pp. 1858– 1866.

[22] D.Harwath,A.Recasens,D.Sur´ıs,G.Chuang,A .Torralba,andJ.Glass,"Jointlydiscoveringvisu alobjectsandspokenwordsfrom raw sensory input," in ECCV, 2018, pp.649–665.

[23] C. Xu, D. Tao, and C. Xu, "Multi-view intact space learning," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37,no. 12, pp. 2531–2544,2015.

[24] F.Wu,X.Jing,J.Zhou,Y.Ji,C.Lan,Q.Huang,an dR.Wang,"Semi- supervised multi-view individual and sharable feature learning for webpage classification," in WWW, 2019, pp.3349–3355.

[25] J. Hu, J. Lu, and Y.-P. Tan, "Sharable and individual multi-view metric learning," IEEE transactions on pattern analysis and machine intelligence, vol. 40, no. 9, pp. 2281–2288,2018.