A Credibility Analysis System for Assessing Information on Social Media

Apoorva M G [1], Dr. Vinay S[2]

[1,2]Department of Computer Science & Engineering, PES College of Engineering, Mandya , Karnataka, India

---***---

Abstract— In   today's competitive environment companies are looking for candidates with strong communication skills and social relations skills. Only candidates with these skills will be a strong team player. But for analysing these skills companies are in lack of time and tool support is needed. In this project we design and implement a tool to analyse the employee job attitude using Facebook. For this we extract the Facebook feature. Analyse the employee attitude among co-workers and do report. Once score received from the attitude analysis module it will give the total ranked score using Pearson correlation to rank the candidates. Also it organizes the scores neatly in a report, so that HR can analyse it easily.

Key Words : *Social Network, Quantification, Attitude analysis*

1. INTRODUCTION

Recruiters look for individuals who would contribute effectively for organizational growth. The importance of a resume is that it acts as the first impression of a candidate. In this competitive world, a candidate should possess a powerful resume that conveys the required information in a manner that it stands out among resumes of contemporaries. The databases of companies hold lakhs of resumes which are unstructured and in free style. The information and the structure contents of resumes will be collection under sub topics; the classification and the representation of information differ from one another.

Hence gathering relevant data from each resume and storing it into the companies' database in a particular format would reduce human effort. There are some difficulties of resume service by unions or commercial companies since they consume too much of time, capacity, money, human effort and so on. These companies require filtered/parsed resumes for the recruitment process. Automated recruitment systems require that Job seekers post their resumes on various websites like Indeed.com, LinkedIn, Naukri.com, Monster.com, Resume builder etc. Certain websites may retrieve unwanted resumes while some may provide very minimum number of resumes. This calls for an approach for qualitative evaluation of resumes.

Individuals act in such manner that would maximize their benefits gained from social interactions Facebook provides an easy to use platform that can be accessed from almost anywhere in the world, to satisfy social needs of people. It can also be used for companies to advertise their products and keeping in touch with their customers. Facebook is also ideal for keeping in touch with large amount of people; a task that was formerly being handled via bulky e-mail message chains.

Virtual community, social networking community, social networking service, online community, are words that are constantly brought up in general conversations, media and business world. Not just aggregates of people, social networks are for sharing social interactions, social ties as well as common space. A virtual community differs from any other community only by being in a "virtual space", it still provides the same sociability support, information and sense of belonging.

Facebook alone has over 955 million active users and over 50% of active users log on to Facebook every day and an average user has approximately 130 friends on Facebook. LinkedIn has over 135 million users but is more of an professional network than casual Social networking has become so popular, that according to Anderson Analytics, 71% of social network users could not live without it. "Happy employees are not productive employees." We hear these conflicting statements made by HR professionals and managers in organizations. There is confusion and debate among practitioners on the topic of employee attitudes and job satisfaction—even at a time increasingly important for organizational success and competitiveness.

2. RELATED WORK

In [1] Keith Douglas McCook ―Organizational perceptions and their relationships to job attitudes, effort, performance, and organizational citizenship behaviours‖ study integrates and expands two models of organization support perceptions, job attitudes, effort, and employee behaviour (i.e., Brown & Leigh, 1996; Netemeyer, Boles, McKee, &McMurrian, 1997). An integrated model was hypothesized, in which Perceived Organizational Support and Perceived Opportunity for Reward impacted job   satisfaction, organizational commitment, and job involvement, which in

turn influenced effort (work intensity and time commitment), which subsequently impacted Organizational Citizenship Behaviours' (OCBs) and in-role performance. Employee – supervisor dyads were surveyed (n = 279), and structural equation modelling was used to test the hypothesized model and several alternative models. Results indicated that the hypothesized model fit the data well, and fit better than several a priori developed alternatives. Inspection of specific parameter estimates indicated that POS and POR impacted job satisfaction, job involvement, and affective organizational commitment. In turn, job satisfaction influenced work intensity, whereas job involvement influenced time commitment. Contrary to predictions, employee effort did not significantly impact in role performance or OCBs. Limitations, contributions, and practical implications are discussed.

In [2] the purpose of this study was to provide insight on attitudes towards Facebook advertising. In order to figure out the attitudes towards Facebook advertising, a snowball survey was executed among Facebook users by spreading a link to the survey. This study was quantitative study but the results of the study were interpreted in qualitative way. This research was executed with the help of factor analysis and cluster analysis, after which Chi-square test was used. This research expected that the result of the survey would lead in to two different groups with negative and positive attitudes. Factor analysis was used to find relations between variables that the survey data generated. The factor analysis resulted in 12 factors that were put in a cluster analysis to find different kinds of groups. Surprisingly the cluster analysis enabled the finding of three groups with different interests and different attitudes towards Facebook advertising. These clusters were analysed and compared. One group was clearly negative, tending to block and avoid advertisements. Second group was with more neutral attitude towards advertising, and more carefree internet using. They did not have blocking software in use and they like to participate in activities more often. The third group had positive attitude towards advertising. The result of this study can be used to help company's better plan their Facebook advertising according to groups. It also
reminds about the complexity of people and their attitudes; not everything suits everybody.

In [3] study identifies three major gaps between HR practice and the scientific research in the area of employee attitudes in general and the most focal employee attitude in particular—job satisfaction: (1) the causes of employee attitudes, (2) the results of positive or negative job satisfaction, and (3) how to measure and influence employee attitudes. Suggestions for practitioners are provided on how to close the gaps in knowledge and for evaluating implemented practices. Future research will likely focus on

greater understanding of personal characteristics, such as emotion, in defining job satisfaction and how employee attitudes influence organizational performance.

In [4] study investigates Facebook users' awareness of privacy issues and perceived benefits and risks of utilizing Facebook. Research found that Facebook is deeply integrated in users' daily lives through specific routines and rituals. Users claimed to understand privacy issues, yet reported uploading large amounts of personal information. Risks to privacy invasion were ascribed more to others than to the self. However, users reporting privacy invasion were more likely to change privacy settings than those merely hearing about others' privacy invasions. Results suggest that this lax attitude may be based on a combination of high gratification, usage patterns, and a psychological mechanism similar to third-person effect. Safer use of social network services would thus require changes in user attitude.

In [5] Real-world applications demand effective methods to estimate the class distribution of a sample. In many domains, this is more productive than seeking individual predictions. At a first glance, the straightforward conclusion could be that this task, recently identified as quantification, is as simple as counting the predictions of a classifier. However, due to natural distribution changes occurring in real-world problems, this solution is unsatisfactory. Moreover, current quantification models based on classifiers present the drawback of being trained with loss functions aimed at classification rather than quantification. Other recent attempts to address this issue suffer certain limitations regarding reliability, measured in terms of classification abilities. This paper presents a learning method that optimizes an alternative metric that combines simultaneously quantification and classification performance. Our proposal offers a new framework that allows the construction of binary quantifiers that are able to accurately estimate the proportion of positives, based on models with reliable classification abilities.

In [6] For man was the first in identifying and naming thequantification problem. A (novel) machine learning task which deals with correctly estimating the number of elements of one class in a set of examples.

Many problems in real applications can be seen as quantification problems. Examples are how many products will be bought? How many clients will be given bank credit? How many pieces will fail? It is especially important when the training dataset does not represent a random sample of the target population. Examples have the same presentation (several input features and a nominal output feature), but
 The test set is considered as a whole versus to apply to a single example alone.

🞏 To the test class distributions versus individual predictions for each example. The output of the quantification problem is a real value, but

🞏 The test set is considered as a whole versus to apply to a single example alone.

The Main features :

🞏 Use a probability estimator instead of a classifier.

🞏 Forman did not consider probabilities because —probability estimates depend explicitly on the class distribution; the calibrated probabilities would become uncalibrated whenever the test class distribution varies".

In [7] the address problem of quantification, a supervised learning task whose goal is, given a class, to estimate the relative frequency (or prevalence) of the class in a dataset of unlabelled items. Quantification has several applications in data and text mining, such as estimating the prevalence of positive reviews in a set of reviews of a given product or estimating the prevalence of a given support issue in a dataset of transcripts of phone calls to tech support. So far, quantification has been addressed by learning a general-purpose classifier, counting the unlabelled items that have been assigned the class, and tuning the obtained counts according to some heuristics. In this article, we depart from the tradition of using general-purpose classifiers and use instead a supervised learning model for structured prediction, capable of generating classifiers directly optimized for the (multivariate and nonlinear) function used for evaluating quantification accuracy. The experiments that we have run on 5,500 binary high-dimensional datasets (averaging more than 14,000 documents each) show that this method is more accurate, more stable, and more efficient than existing state-of-the-art quantification methods.

In [8] Class distribution estimation (quantification) plays an important role in many practical classification problems. Firstly, it is important in order to adapt the classifier to the operational conditions when they differ from those assumed in learning. Additionally, there are some real domains where the quantification task is itself valuable due to the high variability of the class prior probabilities. Our novel quantification approach for two-class problems is based on distributional divergence measures. The mismatch between the test data distribution and validation distributions generated in a fully controlled way is measured by the Hellinger distance in order to estimate the prior probability that minimizes this divergence. Experimental results on several binary classification problems show the benefits of this approach when compared to such approaches as counting the predicted class labels and other methods based on the classifier confusion matrix or on posterior probability

estimations. We also illustrate these techniques as well as their robustness against the base classifier performance (a neural network) with a boar semen quality control setting. Empirical results show that the quantification can be conducted with a mean absolute error lower than 0.008, which seems very promising in this field.

3. METHOLOGY

System Architecture: System architecture is the conceptual design that defines the structure and behaviour of a system. An architecture description is a formal description of a system, organized in a way that supports reasoning about the structural properties of the system. It defines the system components or building blocks and provides a plan from which products can be procured, and systems developed, that will work together to implement the overall system.
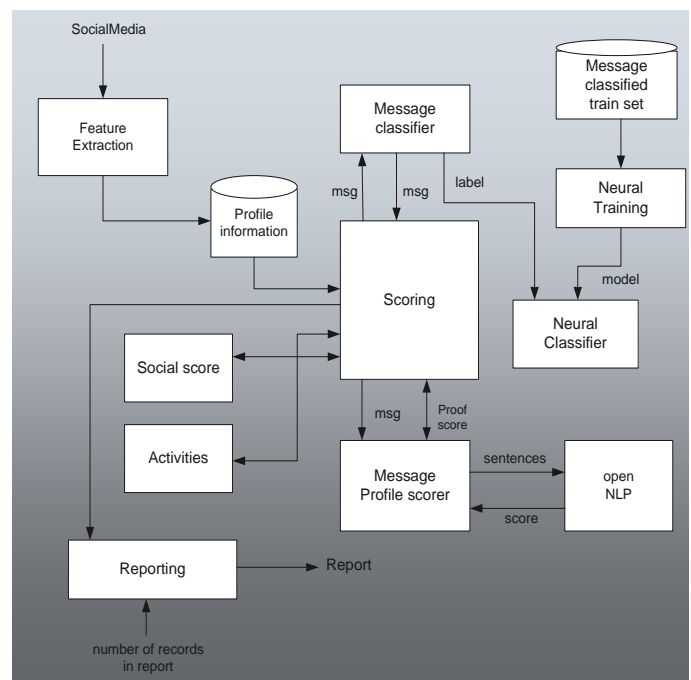
The System architecture is shown above.

In the above architecture diagram each block tells the modules.

Profile Feature Extraction:

This module is implemented using Query Language. It will extract following information from Social Media profile of candidates

It will   following information from Social Media profile of candidates

1. Text messages communicated
2. Number of Friends
3. Groups
4. And interested
5. Games Played and number of times
6. Number of Friends Removed

Attitude Analysis: This module will analyze the following attitudes of candidates from the features gathered from Social Media.

- English Proficiency Level
- Percentage of abusive messages sent
- Number of friends added per month
- Number of groups in profile
- Activity of user in the groups
- Number of Friends removed per month
- Number of friends messaged more times
- Average shared messages in month

Reporting: Once attribute score is received from the Attitude Analysis module, it will give the total ranked score using Pearson correlation to rank the candidates. Also it organized the scores neatly in a report, so that HR can analyse it easily.

4. INTERPRETATION: OF RESULT

The following snapshots define the results or outputs that we will get after step by step execution of all the modules of the system.

INTERPRETATION:

Fig1: new App ID creation
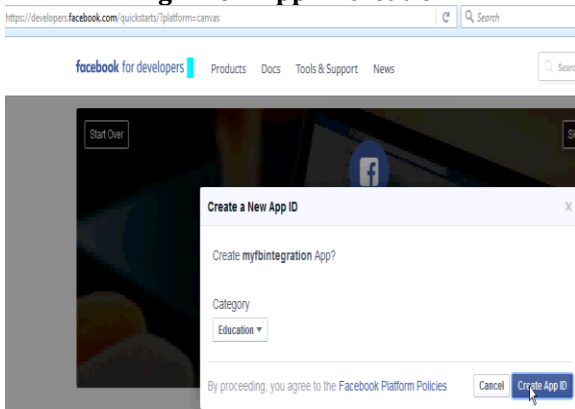
Fig 2: created App id and version with secret key
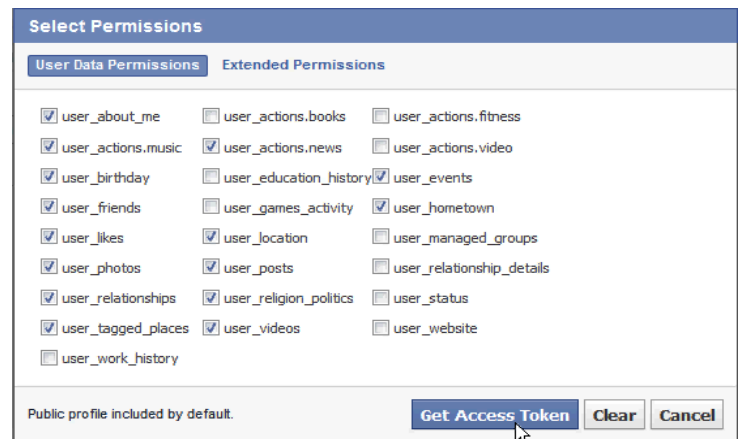
Fig 3: selection of user data permission for accessing token
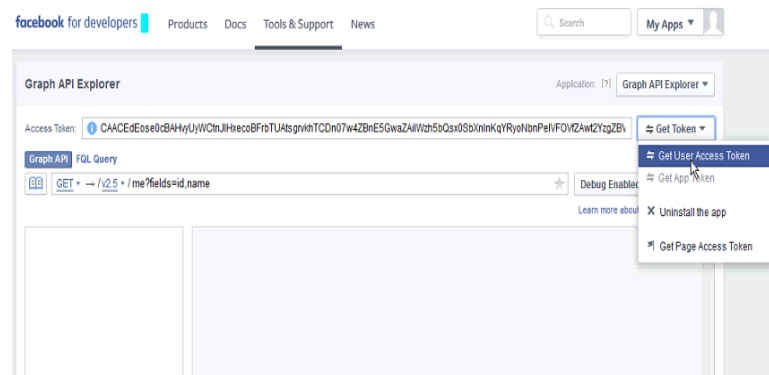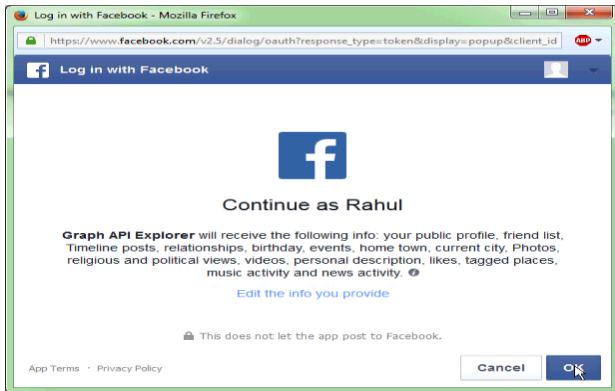

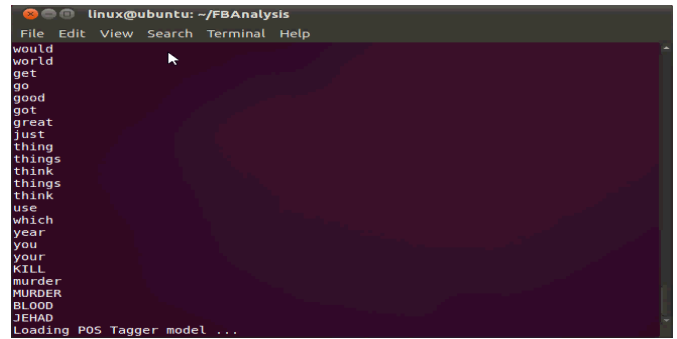**Fig 4: user access token generating**

Fig 5: Facebook account creation

Fig 6: facebook profile information after

Fig 7: running FB analysis in Hadoop platform

Fig 8: loading of points of interest model

Fig 9: normal and abnormal messages vectors

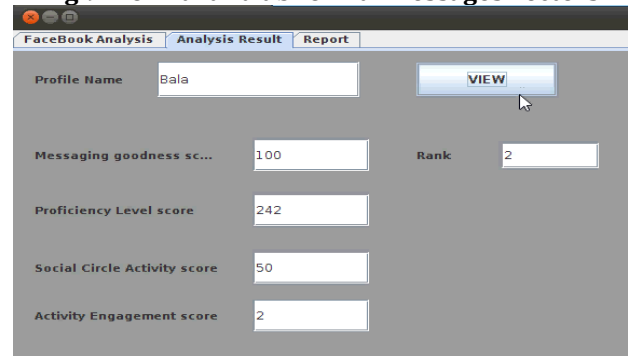**Fig 10: analysing profile history**
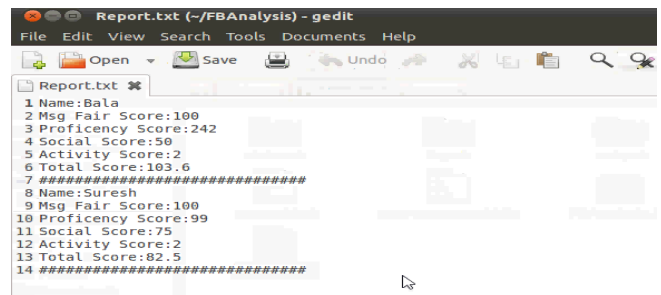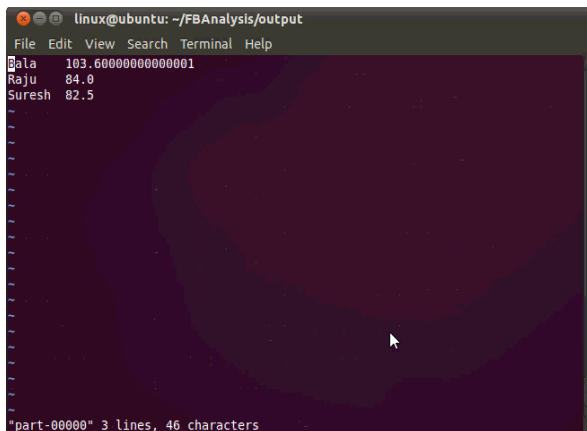
Fig 11: generated user profile report with attributes proficiency score, social score and activity score.
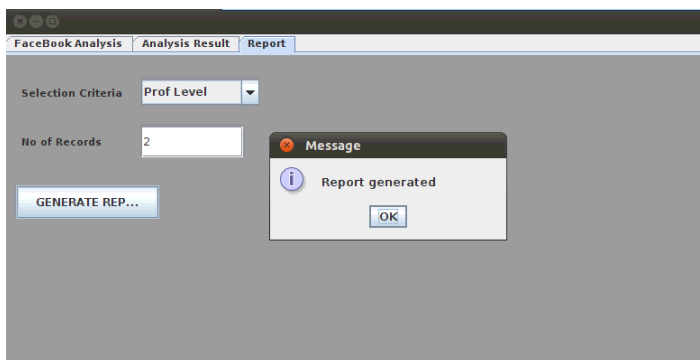
Fig 12: list of profile names

Fig 13: Report generation process based on user profile attributes

5. CONCLUSION

This study utilizes both quantitative and qualitative techniques to explore the timely intersection between online social networking use and privacy concerns. It shows that the gratifications of using Facebook tend to outweigh the perceived threats to privacy. In addition, ongoing research will provide more in-depth understanding of the effects of employee attitudes and job satisfaction on organizational measures, such as customer satisfaction and financial measures.

Future scope

In the near future we will try to analyse the credibility using time-sensitive and location-based approaches that give more reliable and trusted results.

6. REFERENCES

[1]. Keith Douglas McCook "Organizational perceptions and their relationships to job attitudes, effort, performance, and organizational citizenship behaviours" A Dissertation Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College. University of Texas at Dallas, 1997 M. A., Louisiana State University, 1999 December, 2002

[2]. Ville lukka "Attitudes toward Facebook advertising" Survey Paper form Journal of Management and Marketing Research.

[3]. Lise M. Saari and Timothy A. Judge "Employee attitudes and job Satisfaction" Published online in Wiley Inter Science (www.interscience.wiley.com) Human Resource Management ,Winter 2004, Vol. 43, No. 4, Pp. 395–407.

[4] Bernhard Debatin, Jennette P. Lovejoy "Facebook and Online Privacy: Attitudes, Behaviours, and Unintended Consequences" Journal of Computer-Mediated Communication 15(2009) 83– 108©2009 International Communication Association.

[5] Jose Barranquero, Jorge D ıez, and Juan Josedel Coz. Quantification-oriented learning based on reliable classifiers. Pattern Recognition, 48(2):591–604, 2015

[6] Antonio Bella, C`esar Ferri, Jose Hern´andez-Orallo, and Mar´ıa Jos´e Ram´ırez-Quintana. Quantification via probability estimators. In Proceedings of the 11th IEEE International Conference on Data Mining (ICDM 2010)., pages 737–742, Sydney, AU, 2010

[7] Andrea Esuli and FabrizioSebastiani. Optimizing text quantifiers for multivariate loss functions. ACM Transactions on Knowledge Discovery and Data, 9(4):1–27, 2015.

[8] victor Gonzalez-Castro, Rocio Alaiz-Rodrıguez, and Enrique Alegre. Class distribution estimation based on the Hellinger distance. Information Sciences, 218:146–164, 2013

[9] Daniel J. Hopkins and Gary King. A method of automated nonparametric content analysis for social science. American Journal of Political Science, 54(1):229–247, 2010

[10] Gary King and Ying Lu. Verbal autopsy methods with multiple causes of death. Statistical Science, 23(1):78–91, 2008.

Authors Profile

Ms.Apoorva M G Pursuing Master of Technology in Computer Science and Engineering from G Madegowda Institude of Technology in Bharathi Nagara. She has received her Bachelor's degree in Computer Science and Engineering from PES college of Engineering (PESCE), Mandya.

Dr Vinay S currently working as head of department in Department of Information Science and Engineering, PES College of Engineering, Mandya.