

Students Profiling System Based On Mixture Models

G Jyothi*, V.DurgaPrasadarao**, P.SureshBabu***

*(Department of Computer Science, Kaushik College Of Engg., Visakhapatnam-531163)

** (Asst. Professor, Department of Information Technology, Kaushik College Of Engg., Visakhapatnam-531163)

*** (Professor, Department of Computer Science, Kaushik College Of Engg., Visakhapatnam-531163)

ABSTRACT

In order to analyze the user's interest or the navigation patterns, log files obtained from the web servers, help as a rich source of information. The navigation paths chosen by the visitors and the analysis of the browsing patterns helps the users in many applications which include in guiding the users in obtaining relevant information within a short span and for the business personnel to change/plan the strategies. In this paper, we consider the web navigation pattern followed by the students at the time of webcounselling in EAMCET type exam of Andhra Pradesh and thereby helping the managements of the Engineering colleges to understand the student's needs that forced them in choosing particular college and branch. A GMM model is used to assess the students navigating patterns, and helping the managements/users understand the reason beyond the navigation pattern, thereby helping the other students to have an idea about the potential colleges giving placements and the branches which are mostly choosed by the students.

Keywords:- Webnavigation, Webcounselling, Gaussian Mixture Model, EAMCET

1. INTRODUCTION

Web navigation system helps to understand the users navigation pattern [1][2][3][4]. Web navigation pattern are most widely used in literature users different models[5][6][7]. In this paper, we project a scenario of its utilization for students. The students interested in joining the engineering courses at Andhra Pradesh, have to pass the EAMCET exam and have to register through webcounselling process, mentioning their interested branch of study and the name of the engineering college. The students mostly prefer the branches on their interest and the preference will be given to the colleges together with branches basing on the many criteria's among which placement deems to be an important criteria. At the time of webcounselling, the student is supposed to choose 50 colleges as preferences, and among which the allotment will be made based on different criteria by the convenor. However, in order to facilitate the students in choosing the best alternatives among the different colleges is a challenging task. This paper aims to help in this direction, by calculating the PDF of the user's choice and calculating the maximum

likelihood estimate, to determine the student's choice. The rest of the paper is organized as follows. In Section-2, review of the literature is presented, Section-3 of the paper describes the GMM model, Section-4 of the paper deals with the experimental results and final Section-5, concludes the paper.

2. LATEST REVIEW OF THE WORK IN THIS AREA

Navyadhulipala, Tejaswi, Radharani, Parvathydevi (2012). In their paper focused on creating automatic web directory that can be used to list the websites order by category and sub category.

P.K.Srimani, Srinivas (2011) proposed that web personalization can be done by using Recommender Systems and developed a system. This system is used collaborative filtering or content based filtering techniques.

Hirsh, haym. Basu, Chumki. Davison, Brian D (2010) assessed to understand the users behaviour pattern by utilizing log files.

The recent literature in this area suggests that most of the review in this area is driven to analyze the user's patterns, and not much work is reported to develop and analyze a system, which can be useful for the students in deciding their choices. Hence in this paper, a novel methodology is proposed to assist the students by developing a web navigating system which helps the students to decide their choices at the time of admission into various engineering streams. A Gaussian Mixture Model is utilized for this purpose.

3. GAUSSIAN MIXTURE MODEL

The PDF of the Gaussian mixture model is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{x-\mu}{\sigma}\right)^2}$$

where x denotes the number of customers, μ = mean and σ = standard deviation. The user navigation pattern is modelled using the GMM. The user's navigation pattern is identified by calculating the PDF in each case.

4. EXPERIMENTATION

In order to experiment the data, we have obtained the log file, cleaning of the data is done. The data set used consists of the following fields.

List Log

brn	place	clgcd	clgnm
mec	119	JNKP	JNTU KAKINADA COLLEGE OF PHARMACY
mec	120	JNTK	JNTU COLLEGE OF ENGG. KAKINADA
mec	133	ANCUSF	ANU COLLEGE OF ENGG TECHNOLOGY-SELF FINANCE
mec	124	ANUPSF	ANU COLL OF PHARMACEUTICAL SCIENCES
mec	93	CABP	COLLEGE OF AGRICULTURAL ENGINEERING
mec	165	AUCE	A U COLLEGE OF ENGG. VISAKHAPATNAM
mec	249	AUCP	AU COLLEGE OF ENGG. SCHOOL OF PHARMACY
mec	125	AUCPSF	AU COLLEGE OF ENGG.SCHOOL OF PHARMACY-SELF FINANCE
mec	235	AUEVSF	A U COLLEGE OF ENGG FOR WOMEN
mec	133	JNTV	JNTU COLLEGE OF ENGINEERING VIZIANAGARAM
civil	208	JNTH	JNTU COLLEGE OF ENGG. HYDERABAD
civil	87	JNTHMB	JNTU-5 YEAR INTEGRATED MBA SELF FINANCE
civil	238	JNTHMT	JNTU-5 YEAR INTEGRATED M.TECH SELF FINANCE
civil	183	OUC	O U COLLEGE OF ENGG. HYDERABAD
civil	156	OUCT	O U COLLEGE OF TECH. HYDERABAD
civil	125	KUCE	K U COLLEGE OF ENGG. KOTHAGUDEM
civil	191	JNKR	JNTU COLLEGE OF ENGINEERING KARIMNAGAR
civil	157	JNTM	JNTU COLLEGE OF ENGG MANTHANI
civil	195	SVHU	SATAVAHANA UNIVERSITY
civil	91	SVHUSF	SATAVAHANA UNIVERSITY - SELF FINANCE
cse	98	PLMU	PALAMUR UNIVERSITY
cse	195	CASR	COLLEGE OF AGRICULTURAL ENGG
cse	227	KUCP	K U COLLEGE OF PHARMACEUTICAL SCIENCES
cse	154	KUWL	KU COLLEGE OF ENGINEERING AND TECHNOLOGY
cse	135	CAMS	COLLEGE OF AGRICULTURAL ENGINEERING

Table1

Register Log

sno	snm	emailid	pwd
325609247	surya	surya@gmail.c	Student
325609248	shankar	shankar@gmai	Student
325609249	avinash	avinash@gmai	Student
325609250	naidu	naidu@gmail.c	Student
325609251	anuradha	anu@gmail.co	Student
325609252	bhanu	bhanu@gmail.	Student
325609253	prabhudas	prabhud@gmai	Student
325609254	jalli	jalli@gmail.co	Student
325609255	bujji	bujji@gmail.cc	Student
325609256	paul	paul@gmail.co	Student
325609257	suvarna	suvarna@gmai	Student
325609258	shanthi	shanthi@gmai	Student
34534	123456	123@abc.com	password
123456	sample	sample@abc.c	password
356176240	prakash	prakash456@h	Student
995786453	srinivas	srinivaspilla@g	Student
83511531	daniel	ravuridanel@y	Student
597669856	parvathy	parvathy.123@	Student
996282211	sabeena	sabeena.grace	Student
436286654	hemanth	hemanthkuma	Student
706356914	sriya	sriya.567@gma	Student
803425868	kalavani	vanikala@gma	Student
195028802	sindhu	sindhu@gmail.	Student
392372343	raghunadh	raghu@yahoo.	Student
325609244	renuka	renuka@yahoc	Student

Table

Selected Log

sno	brn	place	clgcd	clgnm
325609245	mec	107	AIME	AMALAPURAM INST OF MGMT SCI COLL OF ENGG
325609245	it	112	CFSP	COLLEGE OF FOOD SCIENCE AND TECHNOLOGY
325609245	mec	119	JNKP	JNTU KAKINADA COLLEGE OF PHARMACY
325609245	cse	119	SVUC	S V U COLLEGE OF ENGG. TIRUPATHI
325609245	mec	119	JNKP	JNTU KAKINADA COLLEGE OF PHARMACY
325609245	it	119	SPMUSF	SCHOOL OF ENGG. TECH. SPMVV - SELF FINANCE
325609245	eee	119	SSCC	SRINIVASA INSTITUTE OF TECHNOLOGY AND MANG STUDIES
325609245	ece	119	VITT	VAISHNAVI INST OF TECHNOLOGY
325609245	mec	120	JNTK	JNTU COLLEGE OF ENGG. KAKINADA
325609245	mec	107	AIME	AMALAPURAM INST OF MGMT SCI COLL OF ENGG
325609246	mec	107	AIME	AMALAPURAM INST OF MGMT SCI COLL OF ENGG
325609246	it	112	CFSP	COLLEGE OF FOOD SCIENCE AND TECHNOLOGY
325609246	mec	119	JNKP	JNTU KAKINADA COLLEGE OF PHARMACY
325609246	cse	119	SVUC	S V U COLLEGE OF ENGG. TIRUPATHI
325609246	mec	119	JNKP	JNTU KAKINADA COLLEGE OF PHARMACY
325609246	cse	119	SVUC	S V U COLLEGE OF ENGG. TIRUPATHI
325609246	it	119	SPMUSF	SCHOOL OF ENGG. TECH. SPMVV - SELF FINANCE
325609246	eee	119	SSCC	SRINIVASA INSTITUTE OF TECHNOLOGY AND MANG STUDIES
325609246	ece	119	VITT	VAISHNAVI INST OF TECHNOLOGY
325609246	mec	120	JNTK	JNTU COLLEGE OF ENGG. KAKINADA
325609247	mec	107	AIME	AMALAPURAM INST OF MGMT SCI COLL OF ENGG
325609247	mec	107	AIME	AMALAPURAM INST OF MGMT SCI COLL OF ENGG
325609247	mec	107	AIME	AMALAPURAM INST OF MGMT SCI COLL OF ENGG
325609247	it	112	CFSP	COLLEGE OF FOOD SCIENCE AND TECHNOLOGY
325609247	mec	119	JNKP	JNTU KAKINADA COLLEGE OF PHARMACY

Table3

The data is pre-processed and arranged into table.
 Login Page

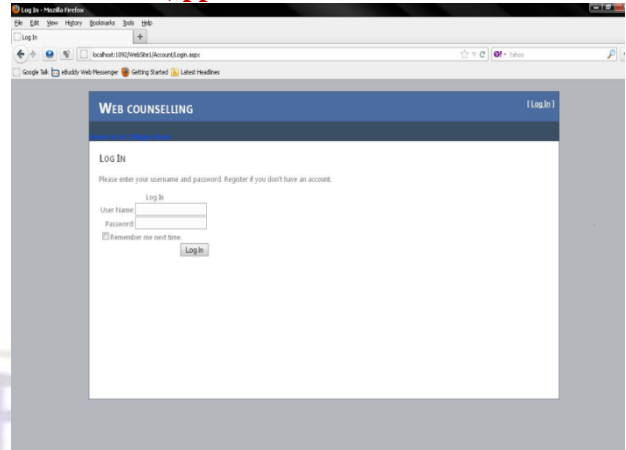


Fig.1 Register Page

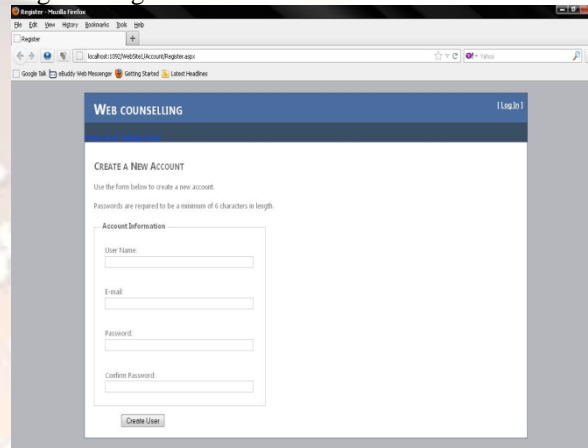


Fig. 2

List of Colleges before sorting

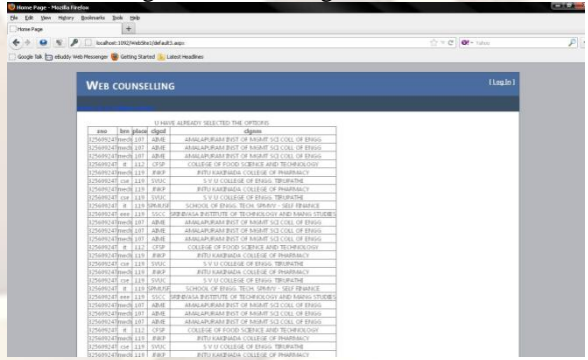


Fig 3

GMM

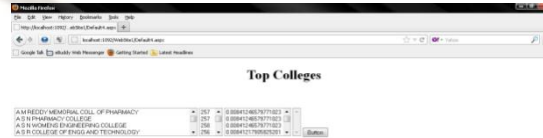


Fig. 4

The experimentation is carried out on PIV system, under dot net environment. The data is first modelled using GMM and the PDF are obtained. The students enter their choices, and to choose the choices, we facilitate a model, using which, the most likelihood estimates (MLE) are obtained for each student choice. The choices are compared with that of the existing preferred choices and the preferences are made.

5. CONCLUSION

In this paper a novel methodology of web navigation using GMM is proposed for benefiting the students in choosing their choices at the time of entering into the engineering stream. The developed model helps to understand the student's pattern and based on the available pattern in the database, the relevant pattern matched is given as output.

REFERENCES

- [1] D. Pierrakos, G. Paliouras, C. Papatheodorou, and C.D.Spyropoulos, "Web Usage Mining as a Tool for Personalization: A Survey," *User Modeling and User-Adapted Interaction*, vol. 13, no. 4, pp. 311-372, 2003.
- [2] J. Srivastava, R. Cooley, M. Deshpande, and P.T. Tan, "Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data," *SIGKDD Explorations*, vol. 1, no. 2, pp. 12-23, 2000.
- [3] Tasawar Hussain, Dr. Sohail Asghar, Dr. Nayyer Masood, "Web Usage Mining: A Survey on Preprocessing of Web Log File", 2010.
- [4] G. Xu, Y. Zhang, and Y. Xun, "Modeling User Behaviour for Web Recommendation Using Ica Model," *Proc. IEEE/WIC/ACM Int'l Conf. Web Intelligence and Intelligent Agent*
- [5] Navya Dhulipala, Tejaswi Pinniboyina, Radha Rani Deevi, Parvathi Devi Karumanchi, "Automated Personalization of Web Directories for usage data using web usage Mining", Jan-Feb 2012.
- [6] P.K.Srimani and Srinivas.A "Website Personalization using Data Mining Techniques-Collaborative Filtering", Vol. 3, Issue, 7, pp.199-203, July, 2011.
- [7] Dimitrios Pierrakos, Georgios Paliouras, "Personalizing Web Directories with the Aid of Web Usage Data", *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, VOL. 22, NO. 9, Sep 2010.
- [8] Xia Min-jie, Zhang Jin-ge 2010. "Research on Personalized Recommendation System for e-Commerce based on Web Log Mining and User Browsing Behaviors," *IEEE international Conference on Computer Application and System Modeling* 12, pp.408-411.
- [9] K R Suneetha and R. Krishnamoorti, 2009. "Identifying User Behavior by Analyzing Web Server Access Log File", *International Journal of Computer Science and Network Security*, 9(4), pp. 327-332.
- [10] Shuqing Wang, and Li She, 2009. "Algorithm Research on User Interests Web Log Data" *Proceedings of IEEE International Conference on Web Information System and Mining*, pp. 93-97.