

# **Prediction of Final Result and Placement of Students using Classification Algorithm**

**Neelam Naik**  
Assistant Professor  
TIMSCDR, Kandivli  
Mumbai, Maharashtra, India

**Seema Purohit**  
Director, NMITD  
Dadar, Mumbai  
Maharashtra, India

## **ABSTRACT**

The quality higher education is required for growth and development of country. Professional education is one of the pillars of higher education. Data mining techniques aim to discover hidden knowledge in existing educational data, predict future trends and use it for betterment of higher educational institutes as well as students. The objective of this study is to use prediction technique using data mining for producing knowledge about students of Masters of Computer Application course before admitting them to the course.

## **General Terms**

Data Mining

## **Keywords**

Data Mining, Higher education, Classification, Prediction

## **1. INTRODUCTION**

Education system is backbone of progressing Indian society. India's superiority in information technology is due to evolution in higher education system in last few years. One has to concentrate on challenges faced by higher educational institutes to improve quality of higher education. Pillania explains that there is lack of research culture in higher education institutes for generating knowledge and lack of interaction between industry and higher education institutes [4]. The industry interaction would help higher education institutes for development of curriculum according to the industry requirements, providing hands on experience for students for applying theoretical knowledge. World Bank report (2010) concentrates on rapidly changing technology needs retraining of workforce [8]. Bhattacharya emphasizes on requirement of repository of latest research in the topics related to India specific needs [2]. The repository should be accessible to students, teachers, researchers, curriculum developers in higher education institutes. According to Rashan, there is a need of improvement in education system by promoting understanding of individual learner capabilities by encouraging problem solving skills, critical thinking and personalized learning [6]. He emphasizes on data mining techniques to gather extra information about academic progress of each student in advance and to help teachers to manage extra coaching and to provide proactive feedback to the student.

Higher Education is categorized into professional and non-professional education. Professional education provides professional knowledge to students so that they can make their stand in corporate sector. Professional education may be technology oriented or it may be totally concentrating on

improving managerial skills of candidate. Masters in Computer Applications (MCA) course provides professional computer technological education to students. This course provides state of the art theoretical as well as practical knowledge related to information technology and make students eligible to stand in progressing information industry.

No system has been developed to predict result of MCA students or to predict placement of student in any company, before admitting him or her in an institute. The prediction of MCA result or the prediction of category of company where student can be placed after the completion of MCA course will help to channelize efforts of students for proper progress. It will also help teachers to take corrective measures towards the progress of the student during the course. It will also help institute management to admit those candidates who will be placed in large scale organizations after completion of the course work. It will help to build reputation of institute in existing similar category institutes in terms of IT education provider and creator of well versatile and skilled employees for IT industry.

The present study concentrates on the prediction of results and placements of MCA students when the student is admitted to the course.

## **2. DATA MINING AS A POWERFUL TOOL**

Data mining is applicable to all sectors of business, like retail, banking, telecommunications, marketing, web mining, medicines and many more. Data mining is widely used in business context but its applications in academics are very much limited. Amar Sahay has developed a software system to assist higher education in assessing and predicting key issues related to student success such as enrollment management, dropout rate, time to degree, and suggests ways to improve courses and programs [1]. Singleton Alex utilized national database, examined systematic grouping of interest within course choice sets and assess whether certain types of student are more likely to make concentrated or diffuse subject selections [7]. This creates an information base of generalized course choice behaviors that higher education institutions could utilize for targeted marketing, recruitment and selection activities, and additionally forms the basis of decision support framework that could be implemented in a variety of online tools to help students. Ranjan Jayanthi et al. propose a framework for an effective educational process using data-mining techniques to uncover the hidden trends and patterns and making accuracy based predictions through a higher level of analytical sophistication in the process of counseling students [5]. Jing Luan addresses the capabilities of data mining and its applications in higher education [3]. Three case studies of J.

Luan demonstrate how data mining saves resources while maximizing efficiency, and increases productivity without increasing cost.

### 3. OBJECTIVE OF THE STDUY

As per the literature review in section 2 earlier works done on applicability of data mining to handle different issues in higher education, but do not focus on the predictability of the placement of the students corresponding to the different parameters that are responsible for making predictions. The enrollment and successful running of professional courses depends upon the placements and the institutes running these courses would like to analyze the students' data to find a concrete association between the parameters affecting the placement activity.

Hence the research objective of the study is to use data mining algorithms to predict final result and placements of graduate student based on previous academic record.

### 4. HYPOTHESIS OF STUDY

Following hypotheses are made and verified through study.

**H1.** The final result of MCA student is always related to the previous academic record of the student such as SSC (Secondary School Certification), HSC (Higher Secondary Certification), Graduation percentage, and Graduation trade.

**H2.** The final placement of MCA student either in large, medium or small scale organization after the completion of MCA course is predictable on the basis of the previous academic record of the student such as SSC, HSC, Graduation percentage, and Graduation trade.

**H3.** MCA students getting marks around 60% at undergraduate, graduate level, receives better employment if his graduation trade is mathematics, physics, information technology or computer science.

**H4.** MCA students from commerce, zoology, chemistry background scores more in MCA examination even if his undergraduate as well as graduate marks are less than or equal to 60%.

**H5.** MCA Students getting marks around 60% at undergraduate as well as graduate level receives better score in MCA

examination, if his trade of graduation is mathematics, physics, information technology or computer science.

### 5. DATA COLLECTION

This study is based on the documents showing previous academic record of students of MCA institutes in Mumbai (India). The sample comprises around 325 students from various MCA institutes in Mumbai. The each record of student contains Gender, SSCE (Secondary School Certificate Examination) percentage, HSCE (Higher Secondary Certificate Examination) percentage, Graduation percentage, Graduation trade, MCA result and details of organization where student has been placed. Around 195 records are used to create model for prediction of result and placement. Remaining 130 records are used to validate developed model. Table 1 shows sample data. XLminer tool is used to apply data mining algorithms and to develop classification tree. Decision rules are developed according to the classification tree generated. These decision rules are implemented in a software using dot net software platform.

### 6. DATA PROCESSING

Data from different sources must be converted into a common format for processing. Some data may be encoded or transformed into most suitable formats. Data reduction may be used to reduce the number of possible data values being considered. Various fields of records are converted into numerical form. Each of the trade of graduation from BCA (Bachelor of Computer Application) to B.Sc.(Bachelor of Science) is numbered from 1 to 7 on the basis of inclusion of IT or computer science related syllabus in graduation curriculum. The gender field values, 0 for male and 1 for female are also included. The output field is percentage of MCA result of every student, which is converted into categorical form and numbered from 1 to 8. The students are placed in various companies like small scale, medium scale or large scale. The categorization of around 30 companies is done on the basis of revenue of company and number of employees working in the company.

Table 1: Sample data showing MCA result of students

no	NAME	SSC	HSC	Gender	GTRADE	GPER	TOTAL	MCA%	MCA%R
1	AWASTHI VARUN C.	73.60	65.33	0.00	4.00	60.57	2360.00	67.43	4.00
2	BAPAT OMKAR S.	80.00	67.83	0.00	9.00	65.43	2666.00	76.17	6.00
3	BATTE DEVANSHREE V.	70.26	64.17	1.00	1.00	64.04	2373.00	67.80	4.00
4	CHOUDHARY PANKAJ S.	66.00	61.00	0.00	3.00	76.45	2394.00	68.40	4.00
5	CHOURASIA RITASHA	61.60	63.17	1.00	3.00	63.68	2120.00	60.57	3.00
6	CHOURASIYA VISHAL B.	69.73	61.83	0.00	5.00	52.13	2220.00	63.43	3.00
7	DESAI MEGHA A.	84.26	80.67	1.00	1.00	74.33	2699.00	77.11	6.00
8	DESHPANDE NEHA S.	72.20	66.80	1.00	2.00	72.25	2361.00	67.46	4.00
9	DIVITY SHRAVAN V.	72.40	62.67	0.00	6.00	64.50	2411.00	68.89	4.00
10	D'SOUZA VERON	61.86	56.33	0.00	3.00	73.80	2221.00	63.46	3.00
11	DUBEY CHANDRASHEKHA	73.60	56.33	0.00	4.00	51.00	2234.00	63.83	3.00
12	GADDAM SRIMANTHKUM	65.06	44.00	0.00	5.00	53.13	2049.00	58.54	2.00
13	GORE RAHUL R.	82.66	69.17	0.00	3.00	60.06	2110.00	60.29	3.00
14	GUDKA KALPESH R.	51.86	56.83	0.00	3.00	77.23	2617.00	74.77	5.00
15	JAIN BHAVIK P.	73.33	53.67	0.00	4.00	54.37	2169.00	61.97	3.00

## 7. DATA ANALYSIS USING DATA MINING

XLMiner 3.2.8 version is used for applying classification algorithm on given data. The generated classification trees for

the prediction of final result and placement of students are shown in figure 1 and figure 2 respectively.

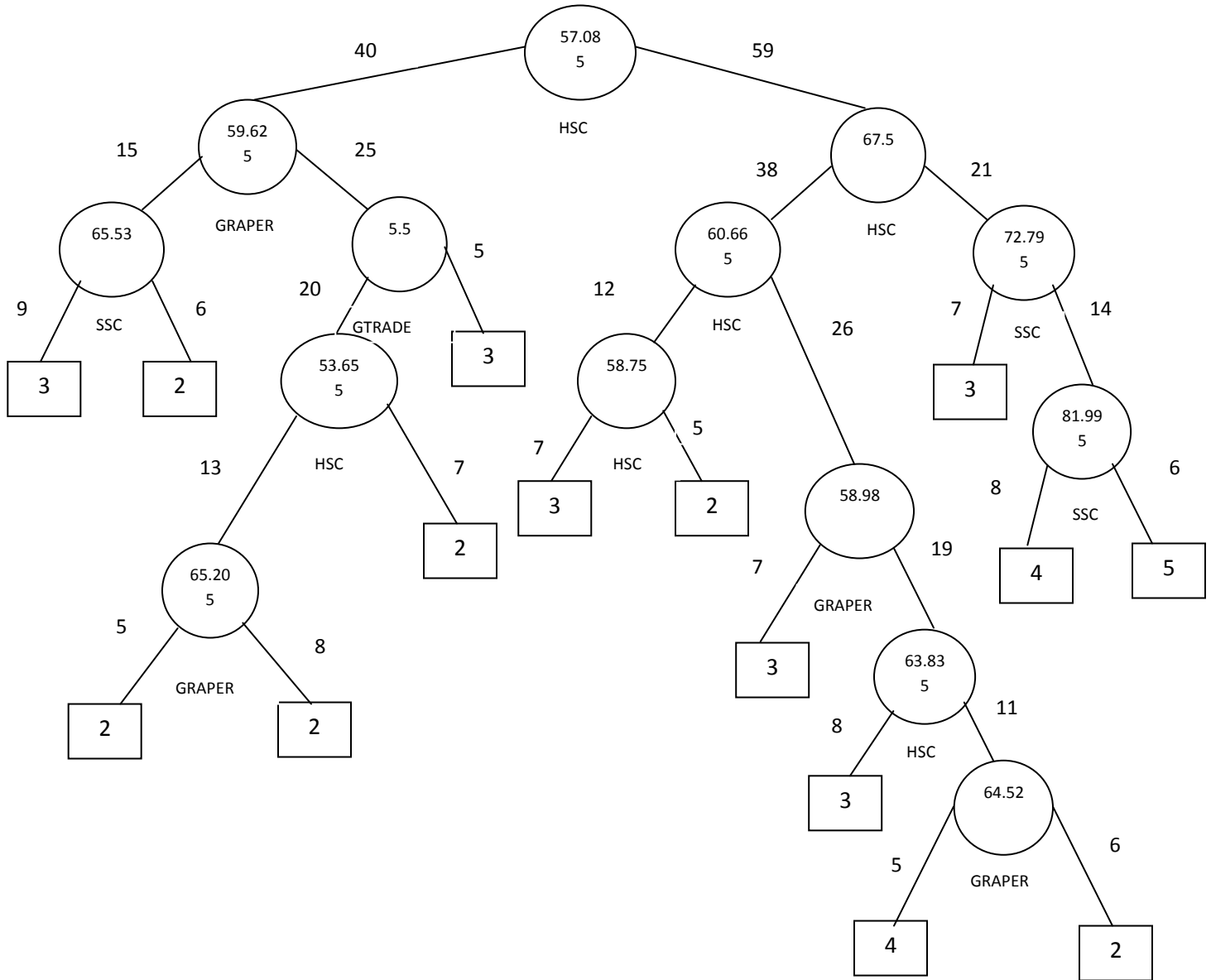
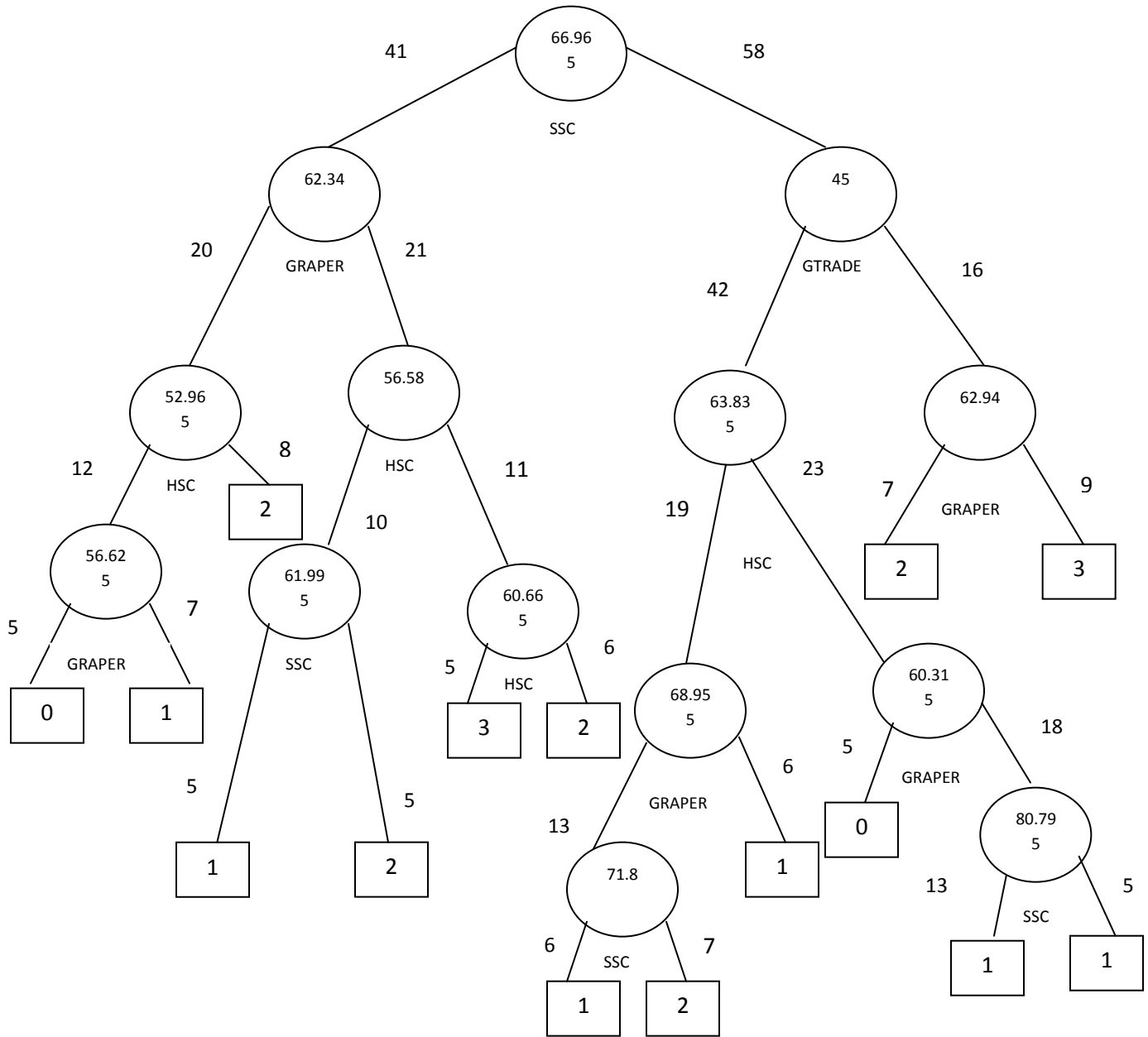


Figure 1: Classification Tree for MCA result- Full Tree (Using Training Data)



**Figure 2: Classification Tree for Placement Prediction - Full Tree (Using Training Data)**

Decision rules are developed based on these trees. ASP.net software is used to develop web forms to implement these rules. Table 2 shows results after testing records on developed tool. In 76.19% cases students (graduation trade: Physics, Mathematics, computer science and information technology) and in 65.38 cases students (graduation trade: Zoology, Chemistry and

commerce ) will secure first class even if during previous academic career student have not secured first class. This supports hypothesis 1 that is the final result of MCA student is always related to the previous academic record of the student such as SSC, HSC, Graduation percentage, and Graduation trade.

**Table 2: Effect of academic record on MCA result and placement of student**

Graduation Trade	SSC Result	HSC Result	GRA Result	MCA Result		Placement	
				< 60%	>= 60%	No	Yes
Physics, Maths, CS, IT	<= 60%	<= 60%	<= 60%	23.80%	76.19%	50.00%	50.00%
Zoology, Chemistry, Commerce	<= 60%	<= 60%	<= 60%	34.61%	65.38%	42.31%	57.69%

Table 2 shows that in 50% cases students (graduation trade: Physics, Mathematics, computer science and information technology) and in 57.69 cases students (graduation trade: Zoology, Chemistry and commerce) will be placed in large scale organization even if during previous academic career student does not have secured first class. This supports hypothesis 2 that is final placement of MCA student either in large, medium or small scale organization after the completion of MCA course is predictable on the basis of the previous academic record of the student such as SSC, HSC, Graduation percentage, and Graduation trade.

If the student is from physics, mathematics, computer science or information technology background and he has secured marks around 60% at undergraduate level as well as graduate level then chances of getting placed in large organization is 50%. This is because as the graduation trade is computer based or mathematical based, they can be technically sound, thorough with their basic concepts and can face placement interview successfully. This supports hypothesis 3 that is MCA students getting marks around 60% at undergraduate, graduate as well as post graduate level, receives better employment if his graduation trade is mathematics, physics, information technology or computer science.

Distinctive features of the research work done as shown in Table 2 indicates that if the student is from commerce, zoology or chemistry background and his under-graduation and graduation percentage is around 60 then the chances of scoring first class at MCA exam is 65%. The reasons for this in the opinion of the researcher are two; one is that as MCA course contains approximately 50% theoretical subjects, students from commerce and non-mathematical science background can score more marks in MCA examination because in theory subjects they can score more than the practical subjects even if they have less marks in their graduation courses and the other is that the mindset of such students is to study hard from the starting of the course as they know that they are from not-IT background.

This supports hypothesis 4 that is MCA students from commerce, zoology, chemistry background scores more in MCA

examination even if his undergraduate as well as graduate marks are less than or equal to 60%.

Table 2 shows that if the student is from physics, mathematics, computer science or information technology background and he has marks around 60% at undergraduate level as well as graduate level then chances of getting first class in MCA examination is 76%. As the graduation trade is computer based or mathematical based, they can score most at MCA level even though their SSC, HSC and Graduation percentage is less than 60. This supports hypothesis 5 that is MCA Students getting marks around 60% at undergraduate as well as graduate level, receives better score in MCA examination, if his trade of graduation is mathematics, physics, information technology or computer science.

## 8. CONCLUSION

The created classification trees for MCA result prediction and placement prediction of students are tested for validation data. The overall error occurred to classify validation data using MCA result prediction classification tree is 38.46% while for validating placement prediction classification tree it is 45.38%. It indicates that final result and placement of students can be predicted before hand admitting them to the course.

The distinctive observation made through present study is that, both the classification trees generated do not take gender as an attribute for taking decision at any level. It indicates that final result and placement prediction of MCA students is not gender specific. For few records it is found that even though academic record of student is not satisfactory and he/she is not from information technology background, placement of such student has been done in large scale organization. The reason behind existence of such outliers is the student's skill to face interview smartly, his/her confidence level and proper behaviour during interview. The knowledge gained through above analysis can be utilized for further enhancement of students; institutes and it can provide feedback for improvement of teaching by faculty members of the higher education institutes.

## **9. REFERENCES**

- [1] Amar Sahay 2010, “Assisting Higher Education in Assessing, Predicting, and Managing Issues Related to Student Success: A Web-based Software using Data Mining and Quality Function Deployment”, Academic and Business Research Institute Conference, Las Vegas
- [2] Bhattacharya 2005. “Changing Strategies in Professional Education” By Prof. S.K. Bhattacharya in IETE technical review, vol. 22, no. 5, September
- [3] Jing Luan 2004, “Data Mining Applications in Higher Education”, SPSS Executive Report
- [4] Pillania Rajesh K. 2007, “Knowledge Management for Indian Business Schools”, Journal of Services Research, Vol. 7, No. 2
- [5] Ranjan Jayanthi and Malik Kamna 2007, “Effective educational process: a data-mining approach”, The journal of knowledge management systems, vol. 37, No. 4
- [6] Rashan K.H. 2010, “Data Mining Applications in the education sector”, MSIT, Carnegie Mellon University
- [7] Singleton Alex 2009, “Data Mining course choice sets and behaviours for target marketing of higher education”, Journal of Targeting, Measurement and Analysis for Marketing Vol. 17, 3, 157–170
- [8] World Bank Report 2010, “<http://www.worldbank.org.in/WBSITE/EXTERNAL/COUNTRIES/SOUTHASIAEXT/INDIAEXTN/0,,contentMDK:21493265~pagePK:141137~piPK:141127~theSitePK:295584,00.html>”, accessed on June 15, 2010