

Result Analysis from PDF File

Akash B. Nikalje

Dept. of Information Technology,
Dr. Vithalrao Vikhe Patil College of Engineering,
Ahmednagar, Maharashtra, India

Ramshekhar V. Benkar

Dept. of Information Technology,
Dr. Vithalrao Vikhe Patil College of Engineering,
Ahmednagar, Maharashtra, India

Amarnath B. Ardhapure

Dept. of Information Technology,
Dr. Vithalrao Vikhe Patil College of Engineering,
Ahmednagar, Maharashtra, India

Shrikisan B. Ghadge

Dept. of Information Technology,
Dr. Vithalrao Vikhe Patil College of Engineering,
Ahmednagar, Maharashtra, India

Prof. Ms. A.B. Bavane

Asstt. Prof., Dept. of Information Technology,
Dr. Vithalrao Vikhe Patil College of Engineering,
Ahmednagar, Maharashtra, India

Abstract— This project paper shows result analysis from pdf file. pdf file is just readable file when the university declare the result then it is hard to analysis this file and categorized this result in the department wise and class wise result of the student. In this paper the exam pdf result calculate and create separate form in first class to fail result and also find pi-chart graph, bar chart and appropriate list of student of subject wise and department wise. This system is developed for analysis of the university result pdf file which includes the student result. Depending on marks of internal exam and final university exam calculate the total marks acquired by student. This gives the statically analyzing of student result and final shows the first class, second class, tact and fail student list of student. As well as mention above show the result subject wise and the in a particular department also this system use java programming language and some data mining concept for developing this application .the adobe portable document file (pdf) format has become a popular means of producing documents for use on other computers when the author cannot be certain of the software available on the other machines. However, once a document is stored in this format, it may generally be considered to be "read only" and the adobe acrobat reader software does not offer a text extraction tool. It discusses third-party software that can translate a pdf to text format.

Keywords-PDF file extraction, data mining , PDF result in college level

I. INTRODUCTION

The introduction of the project paper are the analysis for educational measurement processes test and result analysis of annual examination and the semester examination which is based on the university exam of student. This system is mainly based on the Database technology and the PDF data extraction analysis based system. In this project paper The examination result PDF file are declare by the university and give to the individual colleges this result file in as single format and very hard to distribute in various department and analysis the file who is topper and failure so this paper are show the result calculate and create separate form in first class to fail result

and Also find pie-chart graph, bar chart and separate result of student. This system is developed for analysis of the student Result analysis[1]. The exam has two parts: internal exam and final university exam. Depending On marks of internal exam and final university exam calculate the marks. This gives the statically analyzing of student result and final shows the first class, second class, Tact and fail student list of student. This system use java for developing this application; this application is useful for storing students' information and annual result In the database and shows the student performance. The main goals of this paper are To sort result list and find out how many student are first class, second class, tact, Fail. Our main aim through this project is to reduce manual work, to provide efficient Way of result handling data. The result report module our project reduces the work of Staff of making entries for each department result manually. Result analysis generate Graph of total students grades generation includes by taking only the input PDF file, in future our project can be done as mobile application. So that student can see their progress report on their mobile[1][2].

II. LITERATURE SURVEY

In "students record analysis and examination result computation algorithm (sraerca)" by Abel u. Osagie, Abu Mallam presented data include general information about students as well as examination results at any or all stages of studentship. Analysis of student data can be used to inform solutions to a wide variety of educational institutions. Challenges like exploring group differences, gender representation, Ethnic representation, and growth over time. However, data management and examination result computation of large population of students can Be rigorous time-consuming and prone to errors. Information management Technology offers a comprehensive portfolio of students academic Record to be available in real-time to students and professionals. That a comprehensive solution for use in many tertiary institutions is not yet common presents a unique

opportunity for researchers to help shape the Future of this technology. We introduce a fortran algorithm for data Analysis and result computation (darc). The algorithm analyzes student Data as well as computes examination results of all students in a department At once. Darc reads particular spreadsheet formats of students Information data, semester scores and previous academic records, after Which it outputs, among others, analysis of students information, comprehensive Result format for students showing grade point averages(gpa) And updated cumulative grade point averages (cgpa), summary of academic Records including log of carry over courses, log of carry over Courses that have been cleared, log of outstanding courses and log of Outstanding courses that have been cleared. Using synthetic record of some students, we validate the reliability of the algorithm by conducting Accuracy tests. The algorithm is flexible and can be suitably modified to Manage and preserve academic and general records of both present and past students in any tertiary institution[2][3].

In “result analysis of choice base credit system” by Shubhangi shankar shinde and Dr. Bhatambrekar S. S given educational measurement processes test and result analysis of annual examination based on the university exam of student. This System is mainly based on the database technology and the choice base Credit based system. In this paper the exam result of each subject is calculated Grade wise. This result analysis is based on new credit based System. This system is developed for analysis of the ms (computer Science) result. The exam has two parts: internal exam and external Exam. Depending on marks of internal exam and external exam calculate The marks. This gives the statically analyzing of student result and Final shows the grade of student. This system use Microsoft asp.net for Developing this application; this application is useful for storing students Information and annual result in the database and shows the student performance. The main goals of this paper are to find out if the grades have Had to decline during the years and to evaluate, reliability power of the Test[3][4].

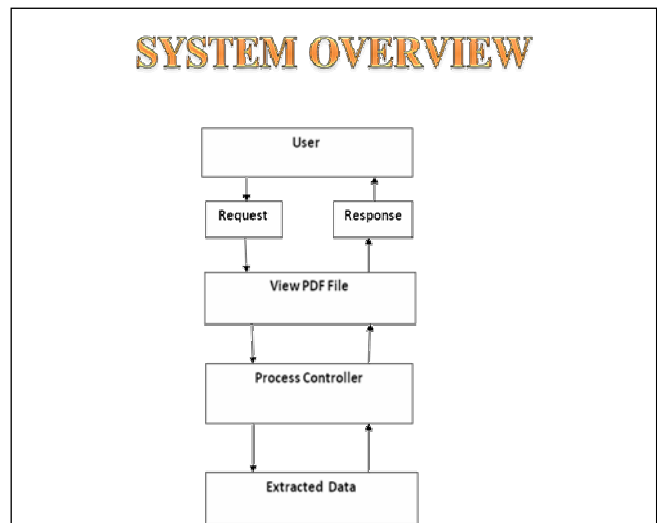
In “online exam cell and result analysis automation” by Aditya Rao, Abhishek Ganesh and Stuti Ahuja given the system is a new concept which came into existence because Of the large amount of data being on paper and it made analysis of Results a tedious task, apart from the unmanageable amount of data that is generated in an institution from various departments. The automation System is like an intermediary between staff and students, thus easing the Activities of each regarding examination. It is a system that will make the exam cell process much organized. It would require certain crucial Data to be pre-existing, which definitely suggest a dependency on certain other systems, especially ones concerning data acquisition. Like in this case it is automated admission system. The solution, however, will manage a great deal of menial work. This keeps paperwork to its minimum, leading to ease of accountability, reducing confusions and increase in work rate and efficiency. The project will address firstly, access to various Users including students, teacher, exam cell staff and admin. Then Customization based on requirement of college

and finally automated Result analysis and ancillary services[4][5].

In “student progress report, result analysis time table generation” by Ghadge Vrushali, Koli Jyoti, Koli Namrata, Koli Shital given the aim of this project is to create software for computer science and engineering department; purpose of this is to maintain the progress Report of each student, to perform result analysis based on semester Result, and to generate timetable. This software gives permission to the Staff to view their proctor details, result details, allotted subject and time Table. The project is more user-friendly so that anyone has minimum Computer knowledge can access and handle the software without having any complexity[5][6][7].

III. PROPOSED SYSTEM

The proposed system are provide to the institute or college level system to analysis generate the result from the pdf file this pdf file are given by the university pdf file are analysis and distributed in college by exam section. It reduce the work load of a exam section in college and manpower because the result sorting and then forward to every department in college is too lengthy process and time consuming so their application system are provide simple and very easy way to sorting and analysis the result in college exam section main thing of system is the result are showing in visualization format to easy to understand for analysis. The result sort and distributed to each department in college. In this project we have been used external library for the extraction of data from the PDF file[7][8]. Because PDF file data is not like simple text format it is in some encrypted formats that why it needs some tool to read those PDF files[9].



PDF input method:

Input PDF file

Library is gives us line by line each and every text from the PDF.

PDF reader module is combine all no. of data lines in one data structure

Extraction method:

Extract data from PDF using iText library.
Extracted data is sends for process.
Process the complete data in parser manner.
Check each and every line for making data set.
Ignore non related data for result generation.

Storage & representation method:

Store complete useful data into the data structure.
Now form the result set for representation.
Now show results in graphical representation using the jfree chart library.
Save graphical result in PDF form.

IV. MODULES

A. Registration and Login Module:

In the registration module the simple some attribute are use like a first name, last name, password for protection and mobile no and email address etc this is must for the registration. Then the registration process is complete after the registration process the user are authorized to the login the login will be successful after the valid username and password then it allow to the enter the system[10][11].

B. PDF Flie Upload Module:

In this module when the login successful then the PDF file are upload the using Drag and Drop as well as download from the internet and browse from the the in system storage when the find or attach the PDF file tit will show the successful msg pop up on the Screen to you that means the PDF file is Uploaded Successfully[12].

C. PDF Extraction and Storage Module :

It is ready for the Pdf extraction module are use when the pdf file is uploaded then the this itext library are read the pdf file line by line and store the result in the buffer memory to generate the result in required format on user demand the Store complete useful data into the memory to generate the appropriate result. Now form the result set for representation this module done task. Now show results in graphical representation using the jfree chart library. Save graphical result in PDF form[12][13].

D. Result generation Module:

In this module the result are display in three format that are given below and also show in the below figure the results in generated result in graph the result, subject result and department result in the graph the two type of graph are generated bar chart and pie chart etc. And then the browse from the folder which are generated by the automatically[13][14].

The result are show by using the imagers of the generated by the PDF file and also email the particular image of the result to the department and other section this is the one of the features of this application. And it provide the how to find the subject wise result and department wise result to analysis the

result how many student are in distinction to class and fail. This is overall structure and working process of this application[14][15].

V. IMPLEMENTATION RESULTS



Fig.1 Registration process

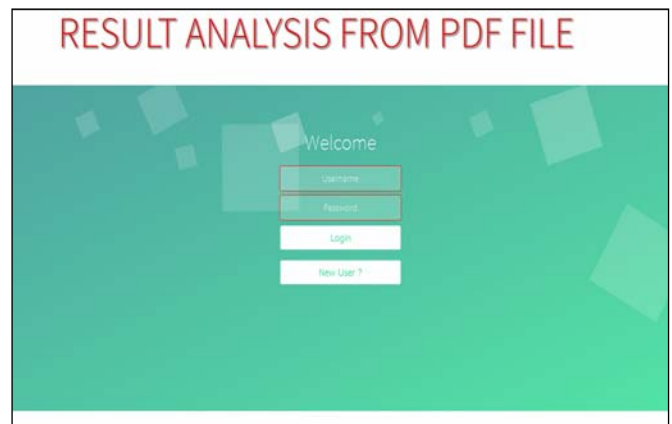


Fig.2 Login after Registration



Fig.3 Home Page



Fig.4 Upload the PDF File

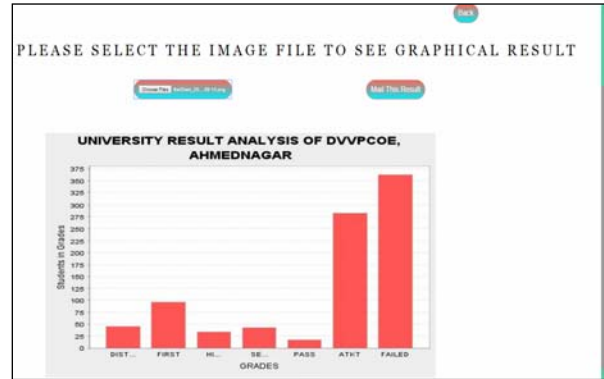


Fig.7 Bar Chart from PDF File

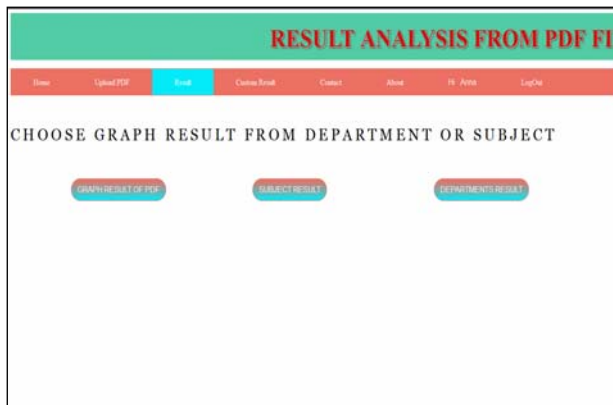


Fig.5 Generating the Result

DEPARTMENTNAME	INFORMATION TECHNOLOGY
GAIKWAD AMRUTA RAJENDRA	1059
KARALE VASUNDHARA DNYANDEY	1050
CHOKHAR PRATIKSHA BALASAHEB	1020
SHELAR KIRAN SHIVAJI	1008
TANTAK ANUSAYA RAJKUMAR	994
NIKALJE AKASH BHIDHARAO	973
SHAIKH ABHIR SALEM	965
BORUDE SWATI DATATRAYA	960
TREDDIKHE SURAJ BALAJI	954
LONDHE PRASHANT RAMDAS	950
DONGARGAVE TERAS RAMESH	949
DANI AKSHAY ABYENO	938
BHOSALE NEHA MAHADEO	925
MEHETRE RUTUJA SANJAY	918
GANDEWAR MONIKA ASHOKRAO	912
ARDHAPURE AMARANTH BALAJI	901
CHITTE HARSHAL BALKRISHNA	900
SUDHEK ARCHANA RAHAJI	900
TODKAL ANIL HARJITI	899
SHEKOLE HARSHADA PANDURANG	888
BERADAR UDAYARAJ SUBHASHIRAO	877
RAMANAND SUNIL KSHIRSAGAR	877
METKAR SANDEEP PRAKASH	872
SINKARANTA BALKRISHNA	861
KHARAT DHANANJAY YASHINTU	856
POTE KANCHAN BALASAHEB	847
PALVE ASHVINI PRAKASH	844
YERIGNTLA TEJAS SANJAY	821
THORAT SANDEEP DNYANESHWAR	813
WAGHMARE SHUBHAM DAMARAO	763

Fig.8 Department Topper Student names

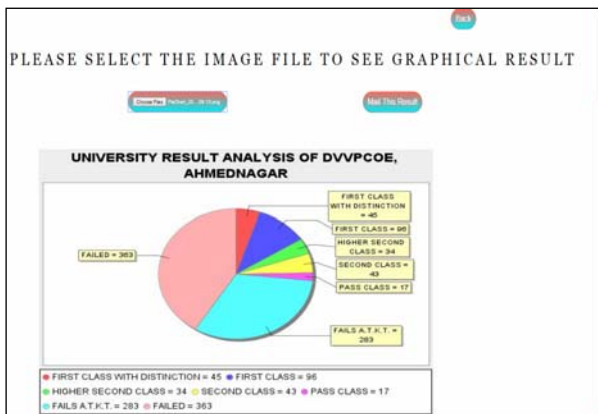


Fig.6 pie Chart from PDF File



Fig.9 Contact Deatils

CONCLUSION

In this paper we implemented system which provides various existing methods used for PDF extraction and how to get the data from the PDF file. The conclusion of this paper is the successfully PDF are extraction and as well as the use the data for our database and generate the appropriate result in various format like pie chart, bar chart and table for the subject wise and department wise result. This paper are give the all simply and easy way to analysis the result of the exam which extracted from the PDF.

ACKNOWLEDGMENT (HEADING 5)

It is our advantage to acknowledge with deep sense of gratitude to our project guide Prof. Ms. A.B. Bavane whose supervision, inspiration and valuable discussion has helped us to complete our project. Their guidance proved to be the most valuable to overcome all the complications in the fulfillment of this project on "Result Analysis from PDF File". We are thankful to Principal Dr. Jayakumar Jayaraman or direct or indirect help in the completion of this project. Last but not least, this acknowledgement would be incomplete without rendering our sincere gratitude to all those who have helped us in the completion of this project.

REFERENCES

- [1] Vidhate, Deepak A and Kulkarni, Parag "Performance enhancement of cooperative learning algorithms by improved decision making for context based application", International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT)IEEE Xplorer, pp 246-252, 2016
- [2] Abel U. O. & Abu Mallam, (2013). "Data Analysis and Result Computation (DARC) Algorithm for Tertiary Institutions", *IOSR Journal of Computer Engineering* (IOSR-JCE, Vol.14, No.3, 63-69. Available:www.iosrjournals.org/iosr-jce/papers/Vol14-issue3/K01436369.pdf?id=7439
- [3] Vidhate, Deepak A and Kulkarni, Parag "Enhancement in Decision Making with Improved Performance by Multiagent Learning Algorithms" IOSR Journal of Computer Engineering, Vol. 1, No. 18, pp 18-25, 2016
- [4] Ukem, E et al. (2012). "A software application for university student's results processing", *Journal of Theoretical and Applied Information Technology*, Vol. 35 No.1. Available: www.jatit.org, (July 2, 2012).
- [5] Vidhate, Deepak, A; Kulkarni, Parag "Multilevel Relationship Algorithm for Association Rule Mining used for Cooperative Learning" in International Journal of Computer Applications, 86(4), pp.20-27, 2014
- [6] Ukem, E. O. and Onoyom-Ita, E. O. (2011). "A software application for the processing of students results", *Global Journal of Pure and Applied Sciences Volume 17* International Journal Of Technology Enhancements And Emerging Engineering Research, Vol 2, Issue 8 Issn 2347-4289, 2014.
- [7] Vidhate, Deepak A and Dhawane, Vandana S "Multimodal Face and Finger Vein Biometric Recognition using ULBP and Repeated Line Tracking" International Journal of Innovations in Engineering and Technology (IJET) vol. 4, No. 3, 2014
- [8] Wayman, J.C., Stringfield, S., &Yakimowski, M. (2004, January). "Software enabling school improvement through analysis of student data", *Center for Research on the Education of Students Placed At Risk (CRESPAR)*, U.S. Department of Education.
- [9] Emmanuel, B. &Choji, D. N. "A software application for colleges of education result processing", *Journal of Information Engineering and Applications, Volume 2, No. 11* (2012).
- [10] Moses E. Ekpenyong. (2008). "A Real-Time IKBS for students results computation", *International Journal of Physical Sciences (Ultra Scientist of Physical Sciences)* Volume 20, Number 3(M),

- September – December, 2008. Available: <http://www.mySQL.com>, (July 22, 2012).
- [11] Isabelle Guyon, Steve Gunn, Asa Ben Hu, Gideon Dror, (2003) "Result Analysis of the NIPS 2003 Feature Selection Challenge", Also Available at :<http://papers.nips.cc/paper/2728-result-analysis-of-the-nips-2003-feature-selection-challenge.pdf>
- [12] Kaspřiková, N. (2012), "Statistical Evaluation of Examination Tests in Mathematics for Economists", *Journal on Efficiency and Responsibility in Education and Science*, vol. 5, no. 4, pp. 203- 211 Also available at <http://dx.doi.org/>
- [13] Somdipdey (2013), New Generation of Digital Academic Transcript using encrypted, Proceedings of IEEE 2013 International Multi Conference on Automation Computing, Communication, Control& Compressed.
- [14] Vidhate, Deepak, A; Kulkarni, Parag (2013) : "Mining Association Rule by Multilevel Relationship Algorithm: An Innovative Approach for Cooperative Learning" in International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), 2(6), pp. 130-137
- [15] Coe, R., (2002). Evidence on the role and impact of performance feedback in schools; In: Visscher, A. J. & Coe, R. (Ed.), *School improvement through performance feedback*, Swetz & Zeitlinger, Lisse, The Netherlands