

DEPARTMENTAL DOCUMENT SOFTWARE USING OCR

**NISHANA SHERIN C ¹, SHAHANA P ¹, RAYHANA BINT YOUSUF ¹, RUSHDA V ¹,
Mr.SHIHABUL HAQ M ²**

¹BSC.Computer Science, KAHM Unity Women's College, Manjeri, Malappuram, Kerala, India

² Assistant Professor , Department Of Computer Science ,KAHM Unity Women's Colleg
Manjeri, Malappuram, Kerala, India

Corresponding author: shihabulhaq2010@hotmail.com

ABSTRACT

This system aims at providing an application for the college departments. The system provides a document processing which aims at entering the university result details by scanning the file and read the contents by using the OCR and prepare an excel sheet automatically, based on that excel sheet the department can analyse the performance of the students based on the auto generated excel sheet and can generate an analysis report of each students, currently the office staff are manually entering the result details of the students into the excel sheets, so this system can reduce the work done by teachers using this document processing.

This system is also used to generate a consolidated report of the attendance of the students based on the daily attendance registration by the teachers. And also the teachers will create an assignment folder in the system so that the students can submit their assignments, based on the submission date and time the teachers can analyse the punctuality of the students, they can get to know who all submitted the assignments on the correct date before due date. By taking the performance analysis of the exam results, attendance and assignments, the system makes it easier for the teachers to calculate the internal marks of the students.

This system also has the provision to enter the fee details of the students according to the community and quota. So that the students can come to know about their scholarships or any other advantages for them if any. And also they can view the college bus fees details, the

number of students using the college bus and any pending fee details etc. The library details of the students can also visible in the system weather they have any fine or any book return or any dues can managed here.

INTRODUCTION

The design and implementation of a College Department Software is to replace the current paper records. The system provides a document processing which aims at entering the university result details by scanning the image formatted file and read the contents by using the OCR and prepare an excel sheet automatically, based on that excelsheet ,the office staff can analyse the performance of the students based on the auto generated excel sheet and can generate an analysis report of each students. Based on department, subjects, year, sem, mark... etc analysis can be done. Office staff is able to access all details of the student. Currently teachers manually enters the result details of the students into the excel sheets, so this system can reduce the work done by teachers using this document processing. The student registration, adding of fee details, university result, file uploading and it's analysis are done by office staff registered by the admin. Admin can register department, office staff, Teacher and Librarian. This system is also used to generate a consolidated report of attendance of the students, based on the daily attendance registration by the teachers. And also the teachers will create an assignment folder in the system so that the students can submit their assignments, based on the submission date and time. The teachers can analyse the punctuality of the students, they can get to know the students who submitted the assignments on the correct date before the due date. By taking the performance of analysis of the exam result, attendance and assignments, the system makes it easier for the teachers to calculate the internal marks of the students. This system also has the provision to enter the fee details of the students according to the admission process (community). The library details of the students are also visible in the system whether they have any fine or any book return or any dues can managed here. Librarian is adding the library details of each student. The details of each student can be viewed by the given teachers, librarian & also office staff in the software.

SYSTEM STUDY

EXISTING SYSTEM

In existing system the office staffs are manually entering the mark details into the system which indeed puts more effort, since they have to enter each student's mark details. Student's

will have to visit the office and department to get to know their results. Teacher also needs to manually add internals and assignment marks into papers and they have to add students attendance into records which takes time and effort. Likewise, librarian also needs to enter the fine and book due details of the student's manually in records which also takes time and effort and which is difficult to sort out a particular student and their fine as well.

Limitation

- Needs lots of time and effort.

PROPOSED SYSTEM

The office staffs have less effort and save time to add marklist because of using Department Software. They can add marklist into the system by using OCR method, which automatically scan the documents and display the image formatted file into Excel sheet. Students do not need to visit office they are likely to know information quickly. Teacher manually adds internals and assignment marks then the system automatically calculates the total mark of the student..when using this technich, we could find the highest mark students per year and also get subject wise top scorers.

METHODOLOGY

The office staffs have less effort and save time to add marklist because of using Department Software. They can add marklist into the system by using OCR method, which automatically scan the documents and display the image formatted file into Excel sheet. Students do not need to visit office they are likely to know information quickly. Teacher manually adds internals and assignment marks then the system automatically calculates the total mark of the student..when using this technich, we could find the highest mark students per year and also get subject wise top scorers.

SYSTEM SPECIFICATION

Hardware and software requirements for the installation and smooth functioning of this product could be configured based on the requirements needed by the component of the operating

environment that works as front-end system here we suggest minimum configuration for the both hardware and software components. Working off with this software is requirements concrete on system environments. It includes two phases.

- Hardware Requirements
- Software Requirements

HARDWARE SPECIFICATION

- Processor : i3 or above.
- System Bus : 32Bit or 64Bit
- RAM :4 GB or Above
- HDD :500 GB or Above
- Monitor :14” LCD or Above
- Key Board : 108 Keys Any Type of
- Mouse :mouse

SOFTWARE SPECIFICATION

- Operating System: Windows 10 Any 32 bit or 64 bit platform
- Front End : Python
- Back End : MySQL Sever
- IDE : Eclipse

SOFTWARE REQUIRED:

FRONTEND:

PYTHON

Python is an interpreted, high-level, general-purpose programming language. Created by Guido van Rossum and first released in 1991, Python has a design philosophy that emphasizes code readability, notably using significant whitespace. It provides constructs that enable clear

programming on both small and large scales. Python is a generalpurpose interpreted, interactive, object- oriented, and high-level programming language. Guido van Rossum during 1985-1990, created it. Like Perl, Python source code is also available under the GNU General Public License (GPL). Python is a popular programming language. Guidovan Rossum created it in 1991.

It is used for:

- web development (server-side),
- software development,
- mathematics,
- System scripting.

Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English keywords frequently where as other languages use punctuation, and it has fewer syntactical constructions than other languages.

- Python is Interpreted – Python is processed at runtime by the interpreter. You do not need to compile your program before executing it. This is similar to PERL and PHP.
- Python is Interactive –You can actually sit at a Python prompt and interact with the interpreter directly to write your programs.
- Python is Object-Oriented – Python supports Object-Oriented style or techniqueof programming that encapsulates code within objects.
- Python is a Beginner's Language – Python is a great language for the beginner-levelprogrammers and supports the development of a wide range of application.

Python's features include:

- Easy-to-learn – Python has few keywords, simple structure, and a clearly definedsyntax. This allows the student to pick up the language quickly.
- Easy-to-read – Python code is more clearly defined and visible to the eyes.
- Easy-to-maintain – Python's source code is fairly easy-to-maintain.
- A broad standard library – Python's bulk of the library is very portable and crossplatform compatible on UNIX, Windows, and Macintosh.
- Interactive Mode – Python has support for an interactive mode which allowsinteractive testing and debugging of snippets of code.
- Portable – Python can run on a wide variety of hardware platforms and has the sameinterface on all platforms.

- Extendable – You can add low-level modules to the Python interpreter. These modules enable programmers to add to or customize their tools to be more efficient.
- Databases – Python provides interfaces to all major commercial databases.
- GUI Programming – Python supports GUI applications that can be created and ported to many system calls, libraries and windows systems, such as WindowsMFC, Macintosh, and the X Window system of Unix.
- Scalable – Python provides a better structure and support for large programs than shell scripting.
-

PHYCHARM IDE

PyCharm is an integrated development environment (IDE) used in computer programming, specifically for the Python language. It is developed by the Czech company JetBrains. It provides code analysis, a graphical debugger, an integrated unit tester, integration with version control systems (VCSes), and supports web development with Django as well as data science with Anaconda.

PyCharm is cross-platform, with Windows, macOS and Linux versions. The Community Edition is released under the Apache License, and there is also Professional Edition with extra features – released under a proprietary license

BOOTSTRAP

Bootstrap is a free and open-source front-end framework for designing websites and web applications. It contains HTML- and CSS-based design templates for typography, forms, buttons, navigation and other interface components, as well as optional JavaScript extensions. Unlike many earlier web frameworks, it concerns itself with end development only. Bootstrap, originally named Twitter Blueprint, was developed by Mark Otto and

Jacob Thornton at Twitter as a framework to encourage consistency across internal tools. Before Bootstrap, various libraries were used for interface development, which led to inconsistencies and a high maintenance burden. After a few months of development by a small group, many developers at Twitter began to contribute to the project as a part of Hack Week, a hackathon-style week for the Twitter development team. It was renamed from Twitter Blueprint to Bootstrap, and released as an open source project on August 19, 2011.

STRUCTURE AND FUNCTION

Bootstrap is modular and consists of a series of Less (Sass version 4 and onward)

stylesheets that implement the various components of the toolkit. These stylesheets are generally compiled into a bundle and included in web pages, but individual components can be included or removed. Bootstrap provides a number of configuration variables that control things such as color and padding of various components. Since Bootstrap 2, the Bootstrap documentation has included a customization wizard which generates a customized version of Bootstrap based on the requested components and various settings. As of Bootstrap 4, SASS is used instead of Less for the stylesheets. Each Bootstrap component consists of an HTML structure, CSS declarations, and in some cases accompanying JavaScript code. Grid system and responsive design comes standard with an 1170-pixel-wide grid layout. Alternatively, the developer can use a variable-width layout. For both cases, the toolkit has four variations to make use of different resolutions and types of devices: mobile phones, portrait and landscape, tablets and PCs with low and high resolution. Each variation adjusts the width of the columns.

BACKEND

My SQL

MySQL is an open source relational database and it includes advanced data types. MySQL operates using client/server architecture in which the server runs on the machine containing the database and client connect to the server over the network. MySQL runs on all platforms supported by MySQL and provides the most direct means of interacting with the server, so it's the logical client to begin with.

- You need to have the MySQL software installed.
- You need a MySQL account so that you can connect to the server.
- You need a database to work with.

The required software includes the MySQL clients and a MySQL client and a MySQL

server. The client program must be located on the machine where you will work. The server can be located on our machine although that is not required.

As long as you have permission to connect to it the server can be located anywhere. In addition to the MySQL software you will need a MySQL account so that the server will allow you to connect and create us sample database and its table.

Microsoft SQL Server 2008 is a full-featured relational database management system (RDBMS) that offers a variety of administrative tools to ease the burdens of database development, maintenance and administration. In this article, we'll cover six of the more frequently used tool: Enterprise Manager, Query analyzer, SQL Profiler, Service Manager, Data Transformation Services and Books Online. Let's take a brief look at each:

Enterprise Manager is the main administrative console for SQL Server installations. It provides you with a graphical "birds-eye" view of all of the SQL Server installation on your network. You can perform high-level administrative functions that affect one or more servers, schedule common maintenance tasks or create and modify the structure of individual databases.

Query Analyzer offers a quick method for performing queries against any of your SQL Server databases. It's a great way to quickly pull information out of a database in response to a user request, test queries before implementing them in other applications, create/modify stored procedures and execute administrative tasks.

SQL Profiler provides a window into the inner workings of your database. You can monitor many different event types and observe database performance in real time. SQL Profiler allows you to capture and replay system "traces" that log various activities. It's a great tool for optimizing databases with performance issues or troubleshooting particular problems.

Service Manager is used to control the MS SQL Server (the main SQL Server process), MSDTC (Microsoft Distributed Transaction Coordinator) and SQL Server Agent processes. An icon for this service Manager to start, stop or pause any one of these services.

Data Transformation Services (DTS) provide an extremely flexible method for importing and exporting data between a Microsoft SQL Server installation and a large variety of other formats. The most commonly used DTS application is the "Import and Export Data" wizard found in the SQL Server program group.

SYSTEM DESIGN

System design is the process of developing specifications for a candidate system that meet the criteria established in the system analysis. Major step in system design is the

preparation of the input forms and the output reports in a form applicable to the user.

The main objective of the system design is to use the package easily by any computer operator. System Design is the creative act of invention, developing new inputs, a database, offline files, method, procedures and output for processing business to meet an organization objective. System design builds information gathered during the system analysis.

The system design is the most creative and challenging phase. The first step is to determine how the output is produced and in what format. Samples of input and output are presented. Next the input data and the master data are to be designed to meet the requirements of the proposed output. The operational phases are handled through program construction testing, including a list of programs needed to meet the system objective and complete documentation.

INPUT DESIGN

Input design is the process of converting the user originated inputs to a computer format. The input design involves determining what the inputs are, how the data should be performed, how to validate data, how to minimize data entry and how to provide a multiuser facility. The design for handling input specifies how data are accepted for computer processing. Input design is a part of overall system design that needs careful attention and includes specifying the means by which actions are taken.

A system user interacting through a system must be able to tell the system whether to accept input produce a report or end processing. The collection of input data is considered to be the most expensive part of the system design. Since the inputs have to be planned in such a manner so as to get the relevant information extreme care is taken to obtain the information. If the data going into the system is incorrect then processing and outputs will magnify this error. All input data are validated in the order and if any data violates any conditions, the user is warned by a message. If the data satisfies all the conditions, then it is transferred to the appropriate tables in the database.

We have to keep in mind the following things to design the system

- What data to input?
- What medium to use?
- The dialogue to guide users in providing input.
- Methods for performing input validation and steps to follow when errors occur

Input requirement gathering was one of the major trivial process in web or android application

development. The project involves text inputs. The inputs can be entered through keyboard and mouse. The text input is gathered by forms with text boxes.

OUTPUT DESIGN

Effective output design will improve the clarity and performance of output. Output design phase of the system is concerned with the convergence of information to the end user friendly manner. The output design should be efficient, intelligible so that system relationship with the end user is improved and thereby enhancing the process of decision making. .

They are also used to provide a permanent copy of these results of processing to the users.

They are also used to provide a permanent copy of these results for late consultation.

There are various types of output required by most systems, the main ones are:

- External outputs, whose destination is outside the organization and which require special attention because they project the image of the organization.
- Internal outputs, whose destination is within the organization and which require careful design because they are the user's main interface with the computer. □
Operational outputs, whose use is purely within the computer department.
- Turn around outputs, to which the data will be added before they are returned to the computer for further processing.

SYSTEM IMPLEMENTATION & MAINTANANCE

The implementation includes all those activities that take place to convert from the old system to new. The old system consists of no filtering the contents searched by the user, which is operated in a push model manner from the proposed new system. A proper implementation is essential to provide a reliable system to meet the requirements of the customers. An improper implementation may affect the success of the application. There are several methods for handling the implementation and the consequent conversion from the old applications to the new application developed in this project. The most secure methods for compare the old system and the new system is to run the old and new system in parallel. In this approach, a person may operate the old existing application and the new application. This method offers high reliability and security. A working version of the system can be implemented in the website application. The website is managed by the admin, Doctor user and the user. The implementation plan

includes host the website and the application put it into its operation. The implementation plan consists of the following steps:

- List all files required for implementation.
- Host the website and put it into its operation.

The implementation plan should anticipate possible problems and must be able to deal with them. The usual problems may be missing documents; mixed data formats between current files and errors in data translation, missing data etc. User Training

The implementation of the proposed system includes the training of system operators. Training the system operators includes not only instructions in how to use the equipment, but also in how to diagnose malfunctions and in what steps to take when they occur. So proper training should be provided to the system operators. No training is complete without familiarizing users with simple system maintenance activities. Since the proposed system is developed in a GUI, training will be comparatively easy than systems developed in a non-GUI. There are different types of training. We can select off-site to give depth knowledge to the system operators.

Success of the system depends on the way in which it is operated and used. Therefore, the quality of training given to the operating person affects the successful implementation of the system. The training must ensure that the person can handle all the possible operations. Training must also include data entry personnel. They must also be given training for the installation of new hardware, terminals, how to power the system, how to power it down, how to detect the malfunctions, how to solve the problem etc. the operators must also be provided with the knowledge of trouble shooting which involves the determination of the cause of the problem.

The proposed system requires trained personnel for operating the system. This will reduce the data entry errors considerably. It is preferable to provide the person with some kind of operating manuals that will explain all the details of the system. For the purpose of training we have improved our user interface for a guiding style of use and we are providing an intuitive interface for users. Along with all the simplicity we are providing a help section for users of the application with a detailed description of how each module is working and feature wise specialties and benefits.

System Maintenance

Maintenance of the software is one of major step in the development of the computer system. Software, which is developed by the engineer, should undergo maintenance process in a regular interval of time as time on new problem arises and it must be corrected accordingly. Maintenance and enhancement are a long- term process.

In this project, the maintenance is carried over by the staff. Since they are the key persons to develop this project they know clearly about the project and coding structured. So, they will change the coding whenever required. Regarding the project maintenance, the changes will occur then and there according to the conditions.

Various types of maintenance that can be made are:

- **Corrective maintenance:** reactive modification (or repairs) of a software product performed after delivery to correct discovered problems. Included in this category is emergency maintenance, which is an unscheduled modification, performed to temporarily keep a software product operational pending corrective maintenance.
- **Adaptive maintenance:** modification of a software product performed after delivery to keep a software product usable in a changed or changing environment. For example, the operating system might be upgraded and some changes to the software may be necessary.
- **Perfective maintenance:** modification of a software product after delivery to provide enhancements for users, improvement of program documentation, and recoding to improve software performance, maintainability, or other software attributes.
- **Preventive Maintenance:** modification of a software product after delivery to detect and correct latent faults in the software product before they become operational faults.

The staff in the concern takes part in each and every level of the project. So they don't need any training of the software. During the development process they sat and entered each and every entry to test the project. They themselves used this is an opportunity to take training is not needed for the users.

CONCLUSION

The project named Department Software is developed as a web application for scanning and displaying the image formatted University marklist into the Excel sheet and its analysis is done at office. This software provides information of a particular student as well. Admin will registers office, teacher and librarian to give information to the particular student to view

his/her academic performance. Students views the assignment, internal and attendance provided by the teacher. And also the student's can know the library fine and book due details provided by the librarian. Thereby each students come to know their information. The project is developed using Python, Django and Sqlite as back end. This language selection is based on the requirement specification of the user and analysis of the existing system, with flexibility for future enhancement. Since the system is developed in modules, future enhancement is very easy.

FUTURE ENHANCEMENT

This Department software is a web application has been designed and developed according to the current techniques and scopes of designing and development tools. This system is veryflexible so that the maintenance and further amendments based on the changing environment and requirements can be made easily with adding further information. Further enhancement is possible to updating evaluating tools. This can be restructured as required. We can do future enhancements mainly on following fields of the product

- Chat: Real time chat between admin and users.
- In further enhancement we aims at introducing text formatted files along with image formatted files . Thus it supports text formatted files to be scanned by OCR method provided. Files can be analysed in many ways based on new features of analysis.
- Office, We will implement updation regarding information given to students on scholarship details and bus fee details rather than fee details.
- The librarian can give information to students about the availability of books, whether the requested book is currently available or not.

REFERENECE

Web References

<http://w3schools.com>, W3Schools

<http://getbootstrap.com>, Bootstrap4

www.wikipedia.com

www.learnvisualstudio.net

[https://www.tutorialspoint.com](https://www.tutorialspoint.com/python/index.htm)

[m/python/index.htm](https://www.tutorialspoint.com/python/index.htm)

www.google.com

