**The Result Shows The Regional Resume Classification System**

**Product Design Book**

|  |  |  |
| --- | --- | --- |
| **File status:**  **[ ]Draft**  **[ ]Under revision**  **[ \* ]Official release** | **Document identification:** | **Product Design Book** |
| **Current version:** | **V1.0** |
| **Document author:** | **Mingyang Duan** |
| **Completion date:** | **2020.04.08** |

|  |  |
| --- | --- |
| **Project team members** | **Students ID** |
| Mingyang Duan | 201930310253 |
| Yue Li | 201930310255 |
| Hengyang Wang | 201930310214 |
| Yiyi Duan | 201930310235 |

**1 Introduction**

**1.1 Product Introduction**

The project team plans to develop a resume classification system, which can be applied to human resource management departments and various talent markets of various types of companies, to help companies and institutions simplify recruitment process and save labor and time costs.

**1.2 Software design objectives and scope**

**Objective:**

To automate resume classification, simplify recruitment process and improve efficiency.

**Scope:**

**Project objective:**

The system can receive text information or extract information from word documents, and display the classification results in a short time .

**Requirements:**

In the face of a large number of resumes, human resource management departments often spend a lot of time and energy on the preliminary screening of resumes. For recruitment, the most important thing is to fully investigate the talents. Therefore, automated resume classification process is a very good measure.

**Restrictions:**

Due to time and technical constraints, the resume text uploaded by users is required to meet the following requirements:

1) Standard format and punctuation;

2) No typos or misspellings;

3) Proper nouns are quoted.

**1.3 Project milestones, project deliverables:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No.** | **Title** | **Date** | **Manager** | **Remarks** |
| 1 | Requirement analysis | 2020.04.03 | Mingyang D |  |
| 2 | Develop project plan | 2020.04.09 | Yiyi D |  |
| 3 | Product design | 2020.04.09 | Mingyang D |  |
| 4 | Project risk plan | 2020.04.12 | Hengyang W |  |
| 5 | Product coding | 2020.04.25 | Mingyang D | Accuracy 70% |
| 6 | Product test plan | 2020.04.26 | Yue L |  |
| 7 | Test code writing | 2020.04.30 | Yue L |  |
| 8 | Product code improvement | 2020.05.03 | Mingyang D | Accuracy 80% |
| 9 | Test code improvements | 2020.05.04 | Yue L |  |
| 10 | Test report | 2020.05.06 | Hengyang W |  |
| 11 | Writing user documentation | 2020.05.07 | Yiyi D |  |

**Milestone Description:**

The project plan is divided into four parts: requirement analysis and plan making, product development and code writing, product improvement and product iteration, product testing.

**The project delivery standards are as follows:**

1) The system can classify at least 5 categories of positions;

2) Each resume is processed within 10s;

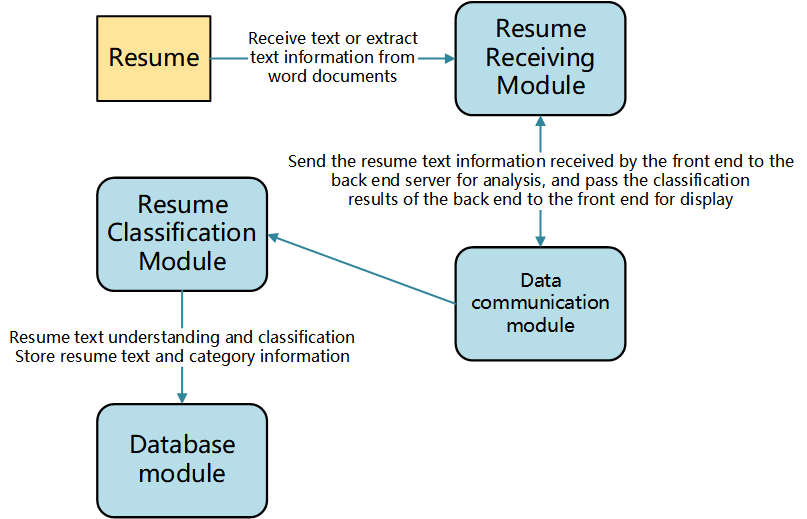
3) The classification accuracy of the system is 75%.

**2 Product overall design**

This section describes the overall design of the system, including functional module design, data flow, database design, front-end page design.

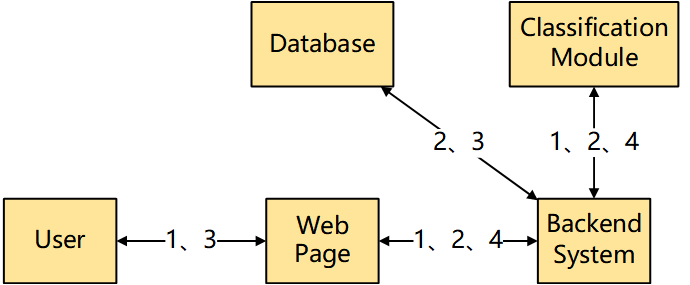
**2.1 Functional module design**

The function module of resume classification system is mainly divided into four parts: resume receiving module, resume classification module, communication module (resume text is transferred from the front end to the back end, and the classification results of the receiving back end are displayed in the front end), database module. The logic of the function module is shown in the following figure:



* Resume receiving module can receive text or extract text information from word document;
* The communication module is used to send the resume text information received by the front-end to the back-end server for analysis, and at the same time, transfer the classification results of the back-end to the front-end for display;
* Resume classification module is used to understand and classify the resume text;
* Database module is used to store resume text and category information.

**2.2 Data flow**



**Data flow description:**

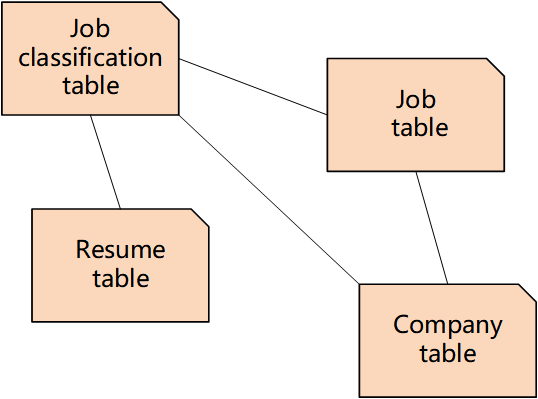
1) User upload resume

2) System classification results

3) Historical information storage

4) Classification result extraction

**2.3 Database design**



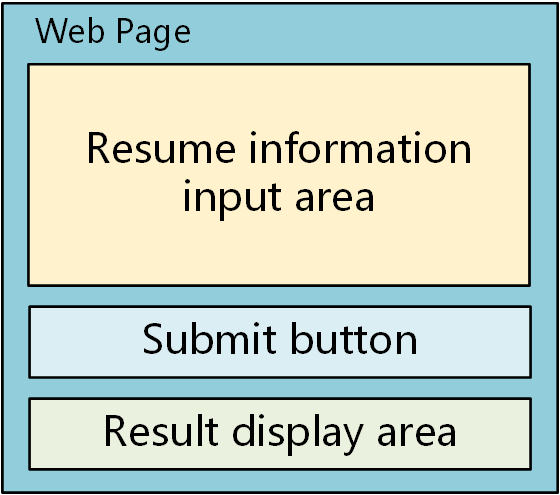
**Database design Description:**

R**esume classification system database contains four tables:**

* + - * The first table is the position classification table. Each field contains four attributes: Company ID, position ID, resume ID and category ID. Resume ID is not only the primary key but also the foreign key;
      * The second table is position table. Each field includes two attributes: Company ID, position ID and position name. The company ID is the foreign key and the position ID is the primary key;
      * The third table is company table. Each field includes two attributes: Company ID and company name. Where company ID is the primary key;
      * The fourth table is resume table, each field includes two attributes: resume ID and resume content. Resume ID is the primary key.

**2.4 front page design**

Besides decoration and frame, the page has three parts: resume input area, submit button and result display area.



**3 Product development technology and methods**

This part will introduce the main technology and framework used in the product development.

**3.1 Front end(Web page)**

HTML, JS, CSS

* HTML is used to write page content
* CSS is used for page layout
* JavaScript is used to write upload image functions and image display functions.

**3.2 Back end(System)**

**Flask**

Flask is a lightweight web application framework written in Python. We use it because of its lightness and strong scalability.

**3.3 Prediction module**

**Keras, Tensorflow**

Keras and TensorFlow are common frameworks for Python deep learning.

**4 Functional module and Interface design**

Functional module and interface design should be the most important part of the project detailed design, which is directly associated to the system implementation.

**4.1 Front end(Web page function)**

**Resume upload;**

**Resume classification;**

**Result display.**

**4.2 Backend(API)**

**Receiving Resume**

**Request method:** POST;

**URL:**Yet to be determined.

**Parameter(s):**

Resume Text(string): utf-8 code

**Response(s):**

Status(int): HTTP Status Code

Result(string): The result of classification

**4.3 Prediction module(Text-CNN)**

**Input:**

**Text (string):** Text read by function Resume\_read()

**Output:**

**result (string):** The result of classification

**4.4 Database module(MySQL)**

The database is used to store and manage resume data and data of various companies and positions, and also to store and manage the relationship between various types of data.