



**İstanbul
Bilgi Üniversitesi**

ENGR 372

PRE-PROCESS ADVISOR FOR
MACHINE LEARNING
ALGORITHMS

BY:

Batıkan Özkul 117200087
Gurur Bilici 11675016
Mahmut Culup 117200012
Burak Seymen 117202076
Ömer Şükrü Varol 118207047

SUPERVISED BY:

METEHAN İNCEGÜL

ISTANBUL BILGI UNIVERSITY

Contents

1	Members and Roles	2
2	Topic of The Project	2
3	Technologies	2
4	Timeline	3
5	Part of the website	3
5.1	Introduction Page	3
5.2	Login Page	4
5.3	Register Page	4
5.4	Myfile Page	5
6	Website Pages	5
7	Future Work	6
7.1	Research And Learning	6
7.2	Design	6
7.3	Implementation	6

June 7, 2021

1 Members and Roles

Gurur Bilici- Backend Developer

Batikan Özkul- Backend Developer

Mahmut Culup- Project Manager, Frontend and Backend Developer

Burak Seymen- Frontend Developer

Ömer Varol- Frontend Developer

2 Topic of The Project

Data preprocessing is the whole process of preparing a data set for applying different algorithms such as machine learning algorithms. Data preprocessing is a very important process because the data which is found and obtained on the internet are raw data generally, it means that data are not suitable for application of algorithms. In some projects, data preprocessing may even take longer than the set up and implementation of the determined algorithm model. In this project, a website was developed that can help solve this problem faced by many engineers, especially software engineers. The dataset uploaded by the user on this website gives feedback to the user about which columns are suitable for machine learning algorithms, thanks to the algorithm running at the backend side, and tells the user which parts of this dataset the user should correct.

3 Technologies

1. Apache NetBeans IDE for PHP, HTML, CSS, MySQL

2. XAMPP and MySQL

3. Visual Studio Code for HTML, CSS, JS

4 Timeline

	Week1	Week2	Week3	Week4
March				Determining Project Topic
April		Meeting for Deciding Task distribution on the Project	Meeting for Website Design ideas	Completing Login Register and welcome pages
May	Milestone Presentation	Meeting between Backend Developers	Started to develop backend	Finished backend Development
June	Successful Testings and Group meeting for presentation			

Figure 1: Figure of the Timeline

5 Part of the website

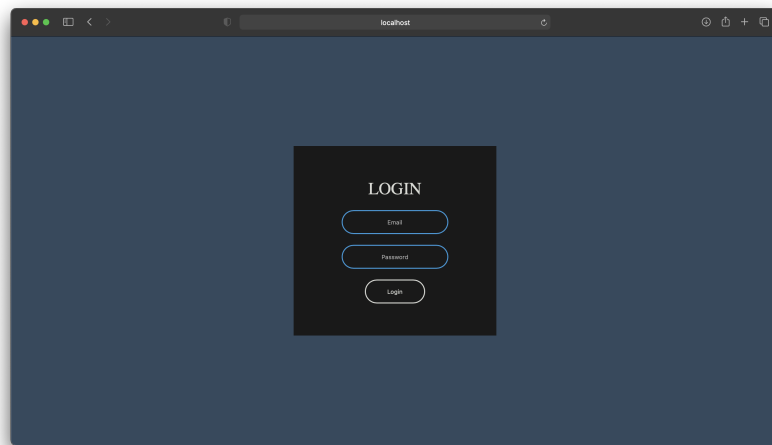
5.1 Introduction Page

- There are two buttons in the top-right side in the page whose names are; Login and Register
- There is a text explaining the importance of data preprocessing, what it is, and information on how to use this website.



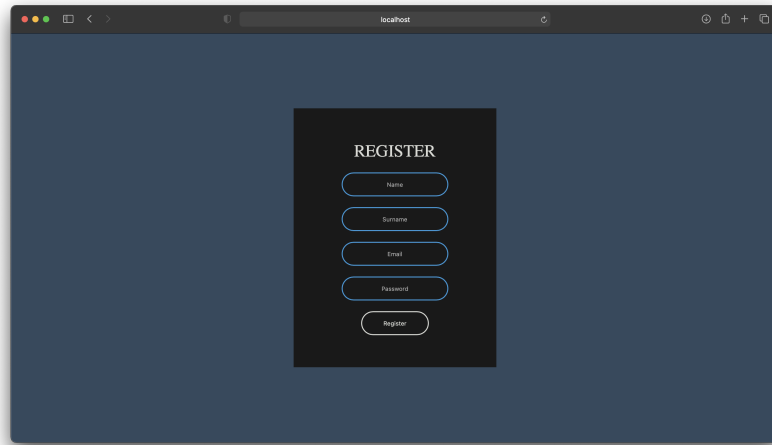
5.2 Login Page

- The login page consists of two text fields;email,password and a login button.
- If the email entered by the user is in the dataset and the password entered by the user and the password corresponding to this email in the dataset match, user can enter the page that user can upload the data and gives feedbacks about the file.
- If the email or password is not in the dataset, the user cannot advance to the next page.



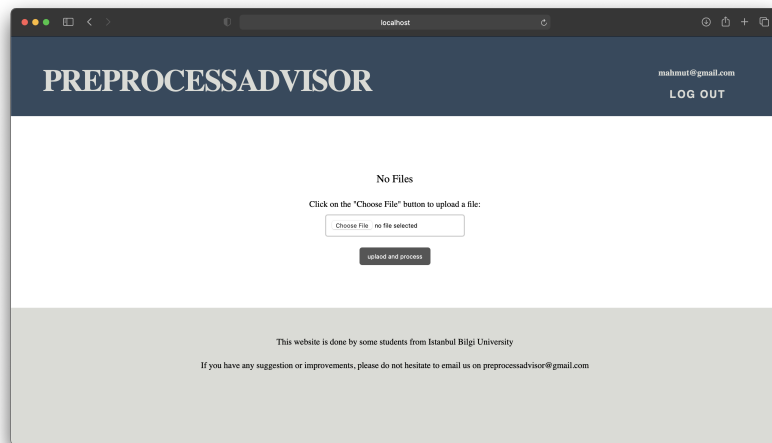
5.3 Register Page

- The Register page consists of four text fields; name, surname, email, password and a button.
- If the e-mail entered by the user is already registered, in the dataset, the user cannot register with this e-mail.



5.4 Myfile Page

- The my files page consists of a "file Upload "section and a button. On this page, the user can see the feedback about the file which is uploaded by user.



6 Website Pages

There are 4 pages of this website. First is the welcome page which contains an explanation about data pre-processing and why is it useful/necessary, it has a straight forward and simplistic design. person who sees it can read it easily and can understand the main idea of the website. Then there is the login and

register pages, they are very basic pages designed by PHP, CSS and HTML. they also contain necessary warning and conditions so for example while registering if someone chooses a invalid e-mail or a short password they will be warned about it. And finally there is the main page where the visitor can use website's services. it contains a table of the data sets user have uploaded before and if he or she did not an empty space. in this table at the right side of the names of the data sets there is a button which calls that items suggestion which was also stored in data set. below this table is the upload button which as you can guess it helps the user to upload files.

7 Future Work

7.1 Research And Learning

- Further improvements on PHP, CSS, SQL, HTML and JavaScript knowledge..

7.2 Design

- Design of Welcome page
- Design of main page

7.3 Implementation

- Python script can be improved and re-implemented in order to evaluate not only csv files but also .jpg and .mp files so the scope of the website can be enlarged.

References

- link to picture1 used in welcome page : "<https://data-flair.training/blogs/python-ml-data-preprocessing/>"
- link to picture2 used in welcome page : "<https://www.google.com/search?q=data+preprocessing&client=dsxsrf=ALeKk02I-dipqJsfJm4q-qVGtaRCgNJqrg:1622967377726&source=lmstbm=ischsa=Xved=wKHZQYC14QAUoAXoECAEQAwbiw=1920bih=955imgrc=AjIMl1MZ-y-nkM>"