M.Sc. (Integrated) Five Years Program
AIML/Data Science
Semester - VII
Department of AIML & Data Science
School of Emerging Science & Technology

SUBJECT CODE: CC-406
SUBJECT NAME: PROJECT –II: PROJECT

MedRec: Medical Report Data Extraction & Maintenance Application

MEMBERS:

SHRUTI HEMANT AGARWAL (DS-01)

PRABLEEN KAUR SANDHU (DS-12)

RADHIKA SHIVKUMAR SHARMA (DS-15)

ESHA MISHRA (AIML-09)

OUTLINE

PROBLEM DEFINITION

For any significant diagnosis or prescription, patient's medical report are a must. On an average a person gets **at least 5 tests done every year**. India is the diabetes capital of the world with 77 million formally diagnosed patients. These people need to get their blood glucose levels regularly tested. Patients with chronic illnesses like thyroid disorder, PCOS/ PCOD get tested every 3 months. Physical file maintenance of these medical reports is laborious and they are easy to misplace or missing when one need them. Maintaining these reports becomes more challenging with time.

OBJECTIVE

The project's main objective is to develop an end-to-end solution for storing and maintaining medical reports using deep learning algorithms and image processing.

INTRODUCTION

80% of all healthcare data is unstructured and inaccessible for further processing. This limits the quantity of usable data and also limits a healthcare organization's decision-making capabilities. According to a study, 30% of healthcare costs are associated with administrative tasks. AI can automate some of these tasks. The global healthcare AI market size is expected to grow from USD 3.64 billion in 2019 to USD 33.42 billion by 2026, at a Compound Annual Growth Rate (CAGR) of 46.21% during the forecast period.

AI-enabled systems are not going to completely replace human medical experts. But this technology will enhance their capabilities and effectiveeness by automating the most repetitive activities prone to errors.

Archiving is the process of securely storing inactive information in any format that you no longer use regularly for long-term retention. Such information is still important to organizations and must be retained for future reference or regulatory compliance.

Medical report storage is a practical solution to free up space while keeping records safe, ensuring they are not lost, stolen, or damaged.

For life-saving documents like medical records, there is no solution available that allows easy access to patient medical history. In case of a medical emergency, easy and fast access to medical history reduces diagnosis time and enables doctors to get the treatment quickly. Big hospitals have their own softwares which will not be accessible to other medical institutions outside the hospital network. Thus, there is a need for a solution that's accessible to all.
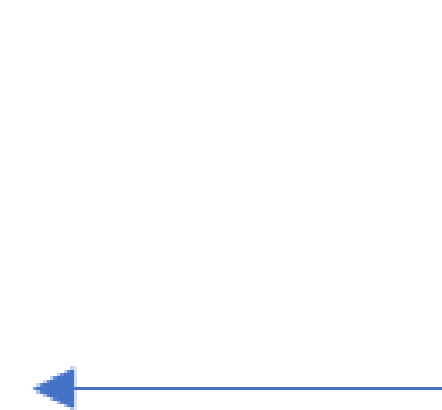
Workflow

Upload Report

Accessing Data

Script / Backend

Display csv file on frontend

Methodology

1   DATA COLLECTION

2   LIBRARIES USED

3   DATA AUGMENTATION

4   DATE EXTRACTION AND PREPROCESSING

5   DATA STORING

6   FRONT END

1

DATA COLLECTION

Collecting Images of medical reports of various tests like blood glucose level, thyroid, urine test, CBC etc. All reports are taken with patient consent and they are aware of the use of these reports.

Report 1

U.N.MEHTA INSTITUTE OF CARDIOLOGY & RESEARCH CENTRE
(Affiliated to B.J Medical College, Ahmedabad)
Civil Hospital Campus, Asarwa, Ahmedabad - 380 016
Phone: 22683911, 22680452 Ext. No.: 113 Fax: 22682092
Visit us at: www.unmicrc.org

DEPARTMENT OF RADIOLOGY

PATIENT NAME: BHRAMADEVI K SHARMA	DATE:27/5/2014
AGE/SEX:76Y/F	OPD.NO:53636/14	RADIOLOGY REGISTRATION No:930
INVESTIGATION: RENAL DOPPLER	

On gray Imaging:

Both Kidneys: Both kidneys are normal in size, shape, axis, show **normal echo pattern** and <u>show mild cortical irregularities</u>.

Approximately 40 x 34 mm size simple cortical exophytic cyst is seen in relation to mid pole of left kidney.

Right kidney- 82 x 32 mm
Left kidney:77 x 40 mm

No e/o calculus or hydronephrosis on either side.
C-M differentiation is well maintained on either side.

On color imaging:

Both intrarenal and extrarenal arteries show normal color filling spectral wave forms and spectral indices.

Both renal arteries at origin appear normal.

IMPRESSSION: NORMAL RENAL DOPPLER STUDY.

Dr.Dinesh L. Patel
RADIOLOGIST
Hon.A.P.of Radiology

Dr. Samir G. Patel
RADIOLOGIST

Dr.Megha Sheth
RADIOLOGIST

Dr. Yashpal R Rana
RADIOLOGIST

Report 2

APEX HEART INSTITUTE ®
(A Unit of TCVS Pvt. Ltd.)

LABORATORY REPORT

Name	: **Mrs. INDRA MISHRA**	Sex/Age	: **Female/ 54 Years**	Case ID	: **20601300571**
Ref.By	: Dr. Tejas Patel MD DM FACC	Dis. At	: OPD	Pt. ID	: 2106151
Bill. Loc.	: Apex Heart Institute OPD			Pt. Loc	:

Reg Date and Time	: 11-Jun-2022 12:49	Sample Type	: Serum	Mobile No	8820280009
Sample Date and Time	: 11-Jun-2022 12:49	Sample Coll. By	: AHIPL	Ref Id1	: 1201064895
Report Date and Time	: 11-Jun-2022 13:55	Acc. Remarks	: Normal	Ref Id2	:

TEST	RESULTS	UNIT	BIOLOGICAL REF RANGE	REMARKS
BIOCHEMICAL INVESTIGATIONS				
Renal Function Test				
Sodium *ISE*	142.00	mmol/L	136 - 145	
Potassium *ISE*	4.70	mEq/L	3.5 - 5.1	
Creatinine *Jaffe, alkaline picrate, kinetic with blank rate correction*	0.70	mg/dL	0.5 - 1.5	
Urea *Urease/GLDH*	30.14	mg/dL	16.6 - 48.5	

Note:(LL-VeryLow,L-Low,H-High,HH-VeryHigh ,A-Abnormal)
* Denotes Test not in NABL Scope

Dr. Rujuta shah
M.B.B.S. D.C.P

Page 5 of 6
Printed On : 11-Jun-2022 14:07
Department of Pathology
Outsourced to Neuberg Supratech Reference Laboratories Pvt. Ltd. (In house lab)
NABL Certificate No. : MC-5010

Block : G-K, Mondeal Business Park, Near Gurudwara, S. G. Highway, Ahmedabad - 380059, Gujarat, India.
E-mail : info@apexheart.in • Website : www.apexheart.in • Mediclaim : 99044 07749 • Emergency : 99044 07755
Appointments / Inquiry : +91-079 4100 5922-25, 99044 07001
CIN : U85110GJ2006PTC047694

Report 3

Triesta Sciences
Precision Diagnostics

HCG adding life to years

LABORATORY REPORT

Name	: DIKSHA MISHRA	Sex/Age	: Female / 24 Years	Case ID	: 20800109479
Ref. By	:	Dis. At	:	Pt. ID	:
Bill. Loc.	: TRIESTA SCIENCEC LAB SOLA	Sample Type	: Whole Blood EDTA	Mobile No.	:
Reg Date and Time	: 06-Aug-2022 15:57	Sample Coll. By	: non NSRL	Ref Id1	:
Sample Date and Time	: 06-Aug-2022 15:57			Ref Id2	:
Report Date and Time	: 06-Aug-2022 21:22	Acc. Remarks	:		

TEST		RESULTS	UNIT	BIOLOGICAL REF RANGE	REMARKS
HAEMATOLOGY INVESTIGATIONS					
HAEMOGLOBIN ELECTROPHORESIS					
By Capillary Electrophoresis					
Hb A (Adult)	L	93.4		96.80 - 97.80	
HbA 2	L	1.4		2.20 - 3.20	
Hb S (Sickle)		00	%		
Foetal Hb	H	5.2		0.00 - 0.50	
Abnormal Haemoglobin		No Other abnormal HB seen.			
Interpretation of Hb Electrophoresis		Mildly raised Hb F with decreased Hb A2. ADV : Clinical correlation, family study and moleculr confirmation are suggested.			

Please note change in reference range of HbA and HbA2.
The Hb electrophoresis is performed by capillary electrophoresis (CE) technique using Sebia system. With this method, charged molecules are separated by their electrophoretic mobility in an alkaline buffer with a specific pH. Separation also occurs according to the electrolyte pH and electro osmotic flow. A high voltage protein separation is then performed and direct detection of the hemoglobins is made at 415 nm at the cathodic end of the capillary.
By using alkaline pH buffer, normal and abnormal (or variant) hemoglobins are detected in 15 separate zones allowing their precise identification. The major zones from cathode to anode are: âA'2 (A2 variant), C, A2/O Arab, E, S, D, G- Philadelphia, F, A, Hope, Bart, J, N-Baltimore and H.
The advantages of CE Sebia are:
- Positive identification & traceability of the specimen through barcodes & absence of transfer of the samples. Interference by plasma is overcome as packed red cells are utilized for analysis.
- Fast simultaneous analysis is done with 08 parallel capillaries within 07 min & overall throughput of 30 samples per hour.
- Fully automatic procedure concerning pre-analytical, analytical and post-analytical steps. Automated procedures of addition of the hemolysing solution to the wells from segment, mixing in the wells & injection without transfer of the hemolysates into the capillaries.
- Direct automatic software selection of normal from abnormal patterns (through a color code). Hemoglobin bands with quantitation and electrophoregrams (curves) are automatically displayed on the screen for final reporting.
- The reproducibility between runs and lots have been shown as < 1.0% for HbA, <2.0% for HbA2 & HbS and <5.0% for HbF.
- Easier identification of the cathodic variants (Hb E, Hb O-Arab and Hb C) & presumptive orientation by identification of the variants in its migration zone. The carbonic anhydrase is not visualized enabling the identification of hemoglobin A2 variants in this migration zone.
- Direct detection provides accurate relative quantification of individual hemoglobin fraction, with particular interest, such as Hb A2 for b-thalassemia trait. In addition, the high resolution of this procedure should allow the identification of hemoglobin variants, in particular, to differentiate Hemoglobin S from D and Hemoglobin E from C. Hemoglobin A2 quantification can also be performed when Hemoglobin E is present.
- Ability to differentiate homozygote from heterozygote subjects, although very low amount of Hb A is better detected than with IEF technique, but lesser than with HPLC. Capillary electrophoresis has been recently adapted to Hb study. CE is the newest FDA-cleared method for the

Note:(LL-VeryLow,L-Low,H-High,HH-VeryHigh ,A-Abnormal)

Page 1 of 3

Dr. Pavan Dave
DCP, DNB (PATH)

Dr. Sandip Shah
M.D. (Path. & Bact.)
Consultant Pathologist

Printed On : 06-Aug-2022 21:27

Laboratory: Triesta Sciences - A Unit of HealthCare Global Enterprises Limited, 3, Sun Chambers, Sola Science City Road, Near Sola Bridge, S G highway, Ahmedabad - 380060, Gujarat.
Lab Helpline 91 79 4041 0166 / 167 Histopathology 91 79 4041 0173 Email: pathology.hccahmd@hcgel.com

Report 1

Registration No : 1201064895
Patient Name : Mrs. INDRA MISHRA
Referred by : SELF

Date : 11/06/2022
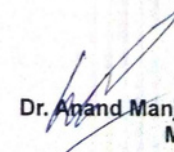Gender/Age : Female / 55 Years

2D Echo WITH COLOUR DOPPLER STUDY

OBSERVATION :

Aortic valve: Trileaflet, sclerosed, no aortic regurgitation.

Other valves are structurally normal.

Normal left ventricle size with Good left ventricular systolic function. (LVEF-60%). No Regional wall motion abnormalities.

Left atrium, Right atrium & Right ventricle are normal in size.

No Pulmonary artery hypertension

Interatrial septum & Interventricular septum are intact.

No clot/vegetation.

No pericardial effusion.

Doppler study revealed no mitral regurgitation, no aortic regurgitation, trivial tricuspid regurgitation.

CONCLUSION :

Normal left ventricle size with Good left ventricular systolic function. (LVEF-60%).

Dr. Anand Manjunath
MD,DM

Report 2

Xpons Diagnostic Centre
(ISO 9001-2008 Certified Lab)

30/11/2, S.M. Ali Road, Kolkata – 700 120.
Phone : Office (033) 2646 0142
e-mail : xponsdiagnosticcentre@gmail.com
West Bengal, Govt. Regd. Log on : www.xpons.in

REPORT

REF. NO. : L-975 (BP/M/DPS)
PATIENT : Ms. BHRAMA DEVI
SEX : F AGE : 76 YRS
Referred By : Dr. S.ADHIKARI MD (CAL)

DATE OF RECEIPT : 09-12-2012
DATE OF REPORT : 09-12-2012

REPORT ON THE BIO-CHEMICAL EXAMINATION

TEST DONE BY SELECTRA PRO XS (FULL AUTO BIOCHEMISTRY ANALYSER) "MERCK"

TEST	TEST VALUE	REF. RANGE
PLASMA GLUCOSE (Fasting)	107.00 mg/dl	(70.00 - 110.00)
SERUM UREA	33.00 mg/dl	(10.00 - 45.00)
CREATININE	0.79 mg/dl	(0.55 - 1.20)

N.B. : All reference ranges are age and sex matched. Reference limits mentioned here are in accordance with the literature provided alongwith the kit which may change with the change in chemistry/method etc.

Dr. S.M. BHATTACHARYA
MBBS (CAL), DCP (CAL)
Consultant Pathologist

Dr. D. BHATTACHARYA
MBBS, MD(Path.)
Consultant Pathologist
Consultant Histopathologist

Help Line : 98314 77468

Report 3

KAMESHWAR MEDICAL CENTRE
(Public Trust Reg. No. E-8329/Ahmedabad)
First Floor, Ankur Commercial Centre, Nr. Ankur Bus Stand,
Naranpura, AHMEDABAD - 380013. M: +91 8141466522
Time : 8-00 a.m. to 5-00 p.m. Website: www.kameshwar.org
☎ 2747 00 87
2748 14 55
7071

Patient's name : ESHA MISHRA
Age/Sex : 17 Years/Female
Referred by :
Receipt No : 23755

Reg. ID : PI-12531-18
Accession : 7071
Order Dt/time : 16/07/2021 08:36
Collection Time : 08:36

Sample Type: Serum

THYROID FUNCTION

Test Name	Result	Units	Biological Reference Interval
Serum T3 Estimation : (Total Tri iodo)	1.19	ng/ml	0.60 to 1.8
Serum T4 Estimation : (Total Thyroxine)	11.7	ug/dl	3.2 to 12.6
TSH(Thyroid Stimulationg Hormone) :	2.25	mIU/L	0.35 to 5.50

DR. BHASKAR SHAH
M.D PATHOLOGIST

2

LIBRARIES USED

- Pandas
- Numpy
- Glob
- Datepasser
- Tensorflow
- Augly
- Opencv
- PIL
- Matplotlib
- Pytesseract
- Datetime
- Pyrebase

3

DATA AUGMENTATION

Data augmentation is a technique used to increase the size and diversity of a dataset by generating new data samples from existing ones. The goal of data augmentation is to improve the performance of a model by providing it with more diverse training data.

```python
# Image path setting
input = 'mr.jpg'
```

```python
input
```

```
'mr.jpg'
```

```python
# Image Scaling with small factor
image = imaugs.scale(input, factor=0.1)
display(image)
```



```python
# Image Scaling with large factor
```

4

DATE EXTRACTION AND PREPROCESSING

To Extract data from image we used OCR (Optical Character Recognition).
Pre-trained model from pytesseract was used to perform OCR on images.

search_dates() from dateparser is used for date extraction.

This extracted data is then processed and invalid dates are filtered out.

```
[ ]  #Extract text from image
     text_dict = pytesseract.image_to_string(img, output_type=Output.DICT)

     print(text_dict.keys())

     dict_keys(['text'])


[ ]  #extrated text
     text_dict['text']

     '  \n \n    \n    \n \n\nAPEX °*\nHEART\nINSTITUTE\n\n(A Unit of TCVS Pvt. Ltd.)\n\n \n\n \n\n   \n  \n\nHe\nHR\n\nLABORATORY REPORT |\n601300571\n\n\n
     Name  : Mrs. INDRA MISHRA Sex/Age : Female/ 54 Years Case ID : 20\n\n         \n\n             \n           \n             \n\nRef.By ; Dr. Tejas Patel MD
     DM FACC Dis. At: OPD pPt.ID  : 2106151\nBill. Loc. : Apex Heart Institute OPD Pt. Loc\n\nReg Date and Time : 11-Jun-2022 12:49 |Sample Type =: S
     erum Mobile No ; 8820280009\nSample Date and Time : 11-Jun-2022 12:49 | Sample Coll. By : AHIPL Refldi  : 1201064895\nReport Date and Time : 11-
     Jun-2022 13:49 |Acc. Remarks _: Normal Ref Id2\n\nTEST RESULTS UNIT BIOLOGICAL REF RANGE REMARKS\n\nBIOCHEMICAL INVESTIGATIONS\n\nLipid Profile
     \n\n  \n\n \n\nCholesterol 136.50 mg/dL 110 - 200\nEnzymatic\n° HDL Cholesterol 47.36 mg/dL 45 - 65\nEnzymatic\nTriglyceride 121.96 mg/dL <150\n
     Enzymatic\nVLDL 24.39 mg/dL 10-40\nCalculated\nChol/HDL 2.88 0 -4.1\nCalculated\nLDL Cholesterol (Direct) 74.31 mg/dL 0.0 - 100.00\nEnzym...'


[ ]  # extracted text  with \n
     ''.join(text_dict['text'].split('\n'))

     '         APEX °*HEARTINSTITUTE(A Unit of TCVS Pvt. Ltd.)      HeHRLABORATORY REPORT |601300571Name  : Mrs. INDRA MISHRA Sex/Age : Female/ 54
     Years Case ID : 20                          Ref.By ; Dr. Tejas Patel MD DM FACC Dis. At: OPD pPt.ID  : 2106151Bill. Loc. : Apex Heart
     Institute OPD Pt. LocReg Date and Time : 11-Jun-2022 12:49 |Sample Type =: Serum Mobile No ; 8820280009Sample Date and Time : 11-Jun-2022 12:49
     | Sample Coll. By : AHIPL Refldi  : 1201064895Report Date and Time : 11-Jun-2022 13:49 |Acc. Remarks _: Normal Ref Id2TEST RESULTS UNIT BIOLOGIC
     AL REF RANGE REMARKSBIOCHEMICAL INVESTIGATIONSLipid Profile   Cholesterol 136.50 mg/dL 110 - 200Enzymatic° HDL Cholesterol 47.36 mg/dL 45 - 65En
     zymaticTriglyceride 121.96 mg/dL <150EnzymaticVLDL 24.39 mg/dL 10-40CalculatedChol/HDL 2.88 0 -4.1CalculatedLDL Cholesterol (Direct) 74.31 mg/dL
     0.0 - 100.00EnzymaticWEW ATP ill GUIDELINES (MAY 2001}. MODIFICATION OF NCEP a oe Oe ups _—TppeHOLESTEROL (| ~—=SGHOLESTEROL = —HDL CHO...'
```

```python
text_dict = pytesseract.image_to_string(img, output_type=Output.DICT)

text = text_dict['text'].split('\n')

from dateparser.search import search_dates
# extrating date fromone record
for i in text:
  if 'date' in i.lower():
    d = i.lower().split('date :')[-1]
    print(d)
    print(search_dates(d))
```

```
patient id unm -2016-01-002376 date: 07-dec-2018
[('07-dec-2018', datetime.datetime(2018, 12, 7, 0, 0))]
```

```python
for name in images:
  print(name)
  record_name.append(name.split('/')[-1])  #storing name of record
  date = date_extractor(name)  # calling extractor
  record_date.append(date)  # saving extracted date
  print(date)
  print('===============================================')
```

```
/content/drive/MyDrive/records/medical records_3.jpg
Extracted text: ['APEX 3%', 'HEART', 'INSTITUTE', '', '(A Unit of TCVS F
All dates found [('Time', datetime.datetime(2022, 12, 18, 0, 0)), ('11-1
final _dates [datetime.datetime(2022, 6, 11, 0, 0), datetime.datetime(20
2022-12-18 00:00:00
===============================================
/content/drive/MyDrive/records/medical records_6.jpg
Extracted text: ['APEX °', 'HEART', 'INSTITUTE', '', '(A Unit of TCVS Pv
All dates found [('Time', datetime.datetime(2022, 12, 18, 0, 0)), ('11-1
final _dates [datetime.datetime(2022, 6, 11, 0, 0), datetime.datetime(20
2022-12-18 00:00:00
===============================================
```

```python
# Date Extraction
def date_extractor(img_name):
  # Read the image
  img = cv2.imread(img_name)
  img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)

  # OCR
  text_dict = pytesseract.image_to_string(img, output_type=Output.

  text = text_dict['text'].split('\n')
  print('Extracted text:', text)

  date = []
  final_dates = []
  date_string_list = []

  #Date Search
  for line in text:
    if 'date'  in line.lower() or 'dt/time' in line.lower():
      date_string_list.append(line)

  date_string = ''.join(date_string_list)
  dates_found = search_dates(date_string)
  print('All dates found',dates_found)
  if dates_found != None:
    final_dates = sorted([d[-1] for d in dates_found  if d[-1] < d
    print('final _dates',final_dates)
    if len(final_dates) >0:
      return final_dates[-1]
    else:
      return 'Not found'

  else:
    return 'Not found'
```

5

DATA STORING

Extracted and cleaned information is converted into csv file format and stored in the database.

For storing data, Firebase is used as it allows us to save data in a much more efficient manner.

```python
import pyrebase
import os


# function for uploading files on  firebase and geting url

def upload(img_path, folder_path='family record/'):

    # firebase configuration details for authentication
    config = {
    "apiKey": "AIzaSyAjL2vfyzO9UKC1_4vuKhy5kkYNyi_rDuQ",
    "authDomain": "medrec-9596f.firebaseapp.com",
    "databaseURL": "https://medrec-9596f-default-rtdb.firebaseio.com",
    "projectId": "medrec-9596f",
    "storageBucket": "medrec-9596f.appspot.com",
    "messagingSenderId": "451422546712",
    "appId": "1:451422546712:web:4d98fd1279043cc6664c6d",
    "measurementId": "G-NPRJEF7YCN"}



    firebase = pyrebase.initialize_app(config)
    storage = firebase.storage()
    img_path = img_path              # where to find the file on drive
    img_name = img_path.split('/')[-1]
    database_path = folder_path + img_name  #where and what name of  file in database

    # Upload Image
    upload = storage.child(database_path).put(img_path)
    # get url
    url = storage.child(database_path).get_url(upload["downloadTokens"])

    return url
```

```python
# storing record names and function in dataframe
data = {'date': record_date,
        'name': record_name}

#creating dataframe
df = pd.DataFrame(data)

#converting date column to datetime type
df['date'] = pd.to_datetime(df['date'], errors='coerce')
# Keepiung only date part not time
df['date'] = df['date'].dt.date
#setting index to date
df = df.set_index('date').sort_values(by=['date'])
```

```python
# dataframe
df
```

	name
date	
2008-01-23	medical_record_8.jpg
2008-06-06	2.png
2012-09-12	medical_record_45.jpg
2014-05-21	medical_record_41.jpg
2014-05-27	medical_record_40.jpg

6

FRONT END

For front end we have build a html web page. We have created an authentication page, home page and display page .

On display Page the extracted csv file is displayed.

```html
<!--Team members-->
<table id="Team Members"  border="0" width="100%" cellpadding="0" cellspacing="0" bgcolor="#CCF2F4">
  <tr>
      <td>
          <table border="0" width="85%" cellpadding="15" cellspacing="0" align="center">
              <!-- Heading Start-->
              <tr>
                  <td height="180" align="center" valign="middle" colspan="4">
                          <font face="arial" color="#000000" size="6">
                              MedRec Team Members
                          </font>
                          <hr width="90" color="#000000">

                  </td>
              </tr>
              <!--Heading End-->
              <tr>

                  <td width="25%" valign="top">
                      <img src="images/e1.jpeg" width="250" height="300" alt="Service1"  width="100%" ></a>
                      <font face="arial" color="#000000">
                          <h3 align="center">Esha Mishra</h3>
                      </font>
                      <font face="arial" color="#426466">
                          Student of M.Sc.Artificial Intelligence and Machine Learning from Gujarat university.
                      </font>
                  </td>
                  <td width="25%" valign="top">
                      <img src="images/s1.jpeg" width="250" height="300" alt="Service1"  width="100%"></a>
                      <font face="arial" color="#000000">
                          <h3 align="center">Shruti Agrawal</h3>
                      </font>
```

Get Started   loginpage.html ✕   index.html

EXPLORER   …

html-practice > loginpage.html > html > body > h1

∨ UNTITLED (WORKSPACE)
  ∨ html-practice
    > images
    index.html
    loginpage.html
    record_details.html

```html
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta http-equiv="X-UA-Compatible" content="IE=edge">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <!-- <link rel="stylesheet" href="loginpage.css"> -->
    <title>Login page in HTML</title>
</head>
<body align="center" bgcolor="#CCF2F4" >
    <h1>Login to Your MedRec Account </h1>
    <form action="" >
        <!-- Headings for the form -->
        <div class="headingsContainer">
            <h3>Sign in</h3>
            <p>Sign in with your username and password</p>
        </div>

        <!-- Main container for all inputs -->
        <div class="mainContainer">
            <!-- Username -->
            <label for="username">Your username</label>
            <input type="text" placeholder="Enter Username" name="username" required>


            <br><br>


            <!-- Password -->
            <label for="pswrd">Your password</label>
            <input type="password" placeholder="Enter Password" name="pswrd" required>

            <br>
```

> OUTLINE
> TIMELINE

⊗ 0 ⚠ 0   Ln 12, Col 18   Spaces: 2   UTF-8   CRLF   {} HTML

Saving File

```
[ ]  #folder where you want to save csv
     save_folder = '/content/drive/MyDrive/csv/'
```

```
[ ]  # saving csv
     df.to_csv(save_folder+'organized_records.csv')
```

```
[ ]  file = pd.read_csv("/content/drive/MyDrive/csvorganized_records.csv")
     file.to_html("record_details.html")
```

OUTPUT

- The output is a csv file with record date, name and url.

- Date helps in organising the documents and url provides easy access to the record image.

- Date extracted is not always accurate and can be improved. For best results the digital image of the document should be clear.

Welcome to MedRec

Your one stop solution for medical records

View your report

Upload

View Report

Login to Your MedRec Account

Sign in

Sign in with your username and password

Your username [Enter Username]

Your password [Enter Password]

[Login]

Not a member? [Register here!](#)

	date	name	url
0	2008-01-23	medical_record_8.jpg	https://firebasestorage.googleapis.com/v0/b/medrec-9596f.appspot.com/o/family%20record%2Fmedical_record_8.jpg?alt=media&token=d08936dd-7d91-4a5d-bb1a-9bb1f2af2f7c
1	2008-06-06	2.png	https://firebasestorage.googleapis.com/v0/b/medrec-9596f.appspot.com/o/family%20record%2F2.png?alt=media&token=d0cd1f90-0179-40c7-bcab-54f60cd959f3
2	2012-09-12	medical_record_45.jpg	https://firebasestorage.googleapis.com/v0/b/medrec-9596f.appspot.com/o/family%20record%2Fmedical_record_45.jpg?alt=media&token=9a30006e-0af4-4c8f-88f2-892987f10f64
3	2014-05-21	medical_record_41.jpg	https://firebasestorage.googleapis.com/v0/b/medrec-9596f.appspot.com/o/family%20record%2Fmedical_record_41.jpg?alt=media&token=fc6b44ee-bf8e-4534-b7a4-75579561d17d
4	2014-05-27	medical_record_40.jpg	https://firebasestorage.googleapis.com/v0/b/medrec-9596f.appspot.com/o/family%20record%2Fmedical_record_40.jpg?alt=media&token=90566f10-e4fa-4c14-a94b-a1ed880217ea
5	2015-02-03	medical_record_43.jpg	https://firebasestorage.googleapis.com/v0/b/medrec-9596f.appspot.com/o/family%20record%2Fmedical_record_43.jpg?alt=media&token=c67924c0-3ae2-425f-bd9a-f616af1b20fa
6	2017-02-16	medical_record_52.jpg	https://firebasestorage.googleapis.com/v0/b/medrec-9596f.appspot.com/o/family%20record%2Fmedical_record_52.jpg?alt=media&token=28a734d1-440d-4ca4-b44d-0fa7be4fc1e8
7	2017-07-09	medical_record_53.jpg	https://firebasestorage.googleapis.com/v0/b/medrec-9596f.appspot.com/o/family%20record%2Fmedical_record_53.jpg?alt=media&token=81992fa2-18e2-4cbd-b4fa-3516dca1e0dd
8	2018-06-20	medical_record_58.jpg	https://firebasestorage.googleapis.com/v0/b/medrec-9596f.appspot.com/o/family%20record%2Fmedical_record_58.jpg?alt=media&token=708db0cf-d724-4dbf-9425-fea1815d68f7
9	2018-08-18	medical_record_56.jpg	https://firebasestorage.googleapis.com/v0/b/medrec-9596f.appspot.com/o/family%20record%2Fmedical_record_56.jpg?alt=media&token=e5651745-314b-45a3-b342-ce186838456e
10	2018-08-18	medical_record_55.jpg	https://firebasestorage.googleapis.com/v0/b/medrec-9596f.appspot.com/o/family%20record%2Fmedical_record_55.jpg?alt=media&token=10c18b35-32c2-43e0-aba8-d3d1235067c0
11	2018-08-18	medical_record_54.jpg	https://firebasestorage.googleapis.com/v0/b/medrec-9596f.appspot.com/o/family%20record%2Fmedical_record_54.jpg?alt=media&token=58119974-4f01-47d8-ae76-b25f77c1df66
12	2018-08-18	medical_record_32.jpg	https://firebasestorage.googleapis.com/v0/b/medrec-9596f.appspot.com/o/family%20record%2Fmedical_record_32.jpg?alt=media&token=55c66fe5-a4e1-4b4d-9bc0-e26573bf666a
13	2018-12-07	medical_record_39.jpg	https://firebasestorage.googleapis.com/v0/b/medrec-9596f.appspot.com/o/family%20record%2Fmedical_record_39.jpg?alt=media&token=d18421d9-cfa5-4ee4-85ff-26225d15b74e
14	2018-12-11	medical_record_50.jpg	https://firebasestorage.googleapis.com/v0/b/medrec-9596f.appspot.com/o/family%20record%2Fmedical_record_50.jpg?alt=media&token=84018d1b-c7a4-41fa-b1f7-28d55975f2e7

CONCLUSION

Medical records are very important life saving documents. Maintaining them is a laborious task. Archiving is used for long term information storage. It has two parts to it. 1.Storing the information so that it is never lost. 2. Easy access to information. Our solution uses firebase by google for storage. The extracted date from medical record images is used to organize the records and save them in a csv file where users can access the record using the given URL. This will help users to store data at a single platform and thus reduce the storage cost and burden of maintaining data physically.

FUTURE WORK

- Improve accuracy of date extraction function.

- A dynamic system that works in real time.

- Authentication for privacy and security.

- Work on the frontend.

REFERENCES

Bajaj Finserv Health. (2022, March 09). Unified Health Interface: Services, Benefits & Registration. Bajaj Finserv Health. Retrieved December 18, 2022, from https://www.bajajfinservhealth.in/articles/unified-health-interface

Data Archives and Why You Need Them. (n.d.). Cloudian. Retrieved December 18, 2022, from https://cloudian.com/guides/data-backup/data-archive/

Improve patient experience with custom medical records management. (n.d.). OSP Labs. Retrieved December 18, 2022, from https://www.osplabs.com/medical-records-management/

MySirG-official/html-project: HTML only Page. (2020, November 25). GitHub. Retrieved December 18, 2022, from https://github.com/MySirG-official/html-project

Personal health records and patient portals. (n.d.). Mayo Clinic. Retrieved December 18, 2022, from https://www.mayoclinic.org/healthy-lifestyle/consumer-health/in-depth/personal-health-record/art-20047273

thisbe jim/Pyrebase: A simple python wrapper for the Firebase API. (n.d.). GitHub. Retrieved December 18, 2022, from https://github.com/thisbejim/Pyrebase

ANY QUESTIONS?

THANK YOU