DECEPTIFORM AND SPEECHER

J COMPONENT PROJECT REPORT

Fall Semester 2020-21

Submitted by

KSHITIJ DHYANI (18BIT0131)
ISHAAN JAIN (18BIT0300)
PRANAV GARG (18BIT0320)

in partial fulfillment for the award of the degree of

B.Tech
in
Information Technology

VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Vellore-632014, Tamil Nadu, India

Under ***Prof. Sharmila Banu K.***

Natural Language Processing (CSE4022)

School of Information Technology

November, 2020

Important links

PROJECT LINKS -

1) DECEPTIFORM WEBSITE: https://deceptiform.herokuapp.com/
2) GIT HUB REPOSITORY: https://github.com/wimpywarlord/women_techies
3) SPEECHR: https://github.com/ishaan14112000/NLP-PROJECT-SPEECHR

HANDS ON TASKS -

1) GITHUB LINK TO MY HANDS ON : https://github.com/wimpywarlord/nlp_hands_on

TEAMS TASKS -

1) https://github.com/ishaan14112000/NLP-NATGEO-GROUP-TASK

ABSTRACT

The name of our product is deceptiform, let us break it down. The first half of the name of the product comes from the movie transformers and the main villains in the same movie are called Decepticons which are humongous robots with autonomous interstellar technology.

The second half is a simple word “Form” which in its entirety means a method to ease the process of exchange of information via a set of input fields.

Lets us Fuse these two words together in an attempt to explain a product. Deceptiform, in other words, means the automation of filling of cumbersome forms which we often find ourselves stuck with, leading to either just staring at the complexly constructed pieces of papers or paying huge loads of money to some professional to decode the form and help us fill it.

Programmed speech recognition is an innovation which has real time genuine applications. From programmed replying mail administrations, to registry enquiries and computerized conveyance administrations, anyplace where people interface with an independent framework, speech recognition gives the connection among man and machine. This venture applies speech recognition to the field of meeting perusing, mechanizing a significant number of the assignments attempted constantly taker. Utilizing an open source speech recognizer framework, that is JAVA, singular speaker accounts are prepared and timestamped. The transcriptions are then processed and integrated into a meeting browser system, in order to facilitate the quick and easy access of meeting information.

MOTIVATION FOR PROJECT

Lets us take an example to better understand the usability of the product. Let us say that you want to buy a piece of land and in order to do that you cant just simply smash a pile of cash on the desk and go on with your life celebrating the new purchase.

On a more simpler alternative reality, you would first have to hire a lawyer to bring about all the paperwork and then you have to sit down multiple times throughout the week, paying for every sitting to that notorious lawyer, trying to decode that encrypted piece of paper that lawyers have carefully curated to make it unreadable to a normal person in order to promote their business model. Enjoy hassle-free documentation with Deceptiform.

Secondly, if you were in a documentation sector you would easily realise that an agreement or a formal doc could go up to 1000-2000 pages in length.

In the face of such a situation, a legal professional or a documentation personal just simply go skimming through the document since they won't be sitting down reading a whole of 2000 pages for every single customer, this assures a guarantee of atleast missing out on atleast a few details, but not with deceptiform. Enjoy a 100% error-free documentation and form filling in lightening fast terms.

Additionally, we are all familiar with the scams that go around these paper works. There could be an insurance paper amongst the other pile of papers, and the lawyer could make you sign with a blind eye everywhere. The only thing you need to keep a blind eye on is the quality and safety of Deceptiform.

Who is to save us amongst such dismay? Now, what if I told you that there was free service that would automatically fill the whole legal document just based on a few questions.

All you need to do is enter the form template which is quite freely available on the internet and sit back while our service using NLP and certain algorithms, constructs a small questionnaire about the details which it will need from you to fill the form.

Just answer a few super simple straightforward question and witness our product smarty fill up the whole document. BOOM, enjoy your document hassle-free.

Let's not limit this application to legal documents, we come across such documents on very frequent occasions, for instance when you wanted to Join Vellore Institute of Technology – Vellore, you were handed hundreds of papers to fill and a thousand places to sign, or whenever you are looking to join a new organisation as an employee or whenever you end up going to the hospital or the notary.

Be it a birthday invite or a simple non-disclosure agreement, simply enter the form template and enjoy the end product while you sip on your morning tea.

Natural Language Processing NLP refers to the evolving set of computer and AI-based technologies that allow computers to learn, understand, and produce content in human languages. The technology works closely with speech/voice recognition and text recognition engines. While text/character recognition and speech/voice recognition allows computers to input the information, NLP allows making sense of this information. So, in this report we are making our effort to show the use of NLP programs. Java speech recognition is one of the project as we are using Sphinx 4 library.

METHODS AND TECHNOLOGY USED

DECEPTIFORM – Deceptiform is a very simple yet novel implementation of google's text to speech api and microsoft's text summarization api. The interface is a web based application, built upon HTML CSS and Vanilla JS. The architecture can be divided into client side and server side. The client side utilizes technologies such as bootstrap, SCSS, where the server side is built upon technologies such as NODE.js, express.js and mongodb. The api used are the following and can be viewed from the respective links:

1) Text to speech api : https://cloud.google.com/text-to-speech

2) Text summarization api : https://rapidapi.com/textanalysis/api/text-summarization

SPEECHER - The programming language used in the project was Java. The IBM EclipseIDE was used to build and manage the speech recognition package, the CMUSphinx 4 Recognizer. One of the requirements of the program was the installation of ApacheAnt to compile the Sphinx package.
Apache Ant is a program for creatingXML build files for Java, and also integrates fully into the Eclipse IDE.With Sphinx being a complicated package containing many separate files and folders. Components of the platform use the Java Speech API, and this was given the source code. The JSAPI is a Java package for use with speech recognition frameworks and speech blend frameworks, and isn't given the Standard Edition of the Java platform.
Using Eclipse gave a steady platform to fabricate and run the Sphinx project. It was conceivable to alter all info boundaries and a debugger was accessible when things turned out badly. The Sphinx site gave directions on the best way to get everything going utilizing Eclipse.The Sphinx framework was picked due to the way that it was executed in Java, and furthermore on the grounds that it was an open source framework and thus liberated from permitting issues. The adaptable idea of the framework, and the way that it was configurable to suit a scope of various errands likewise demonstrated helpful to the project.

CODE AND OUTPUT

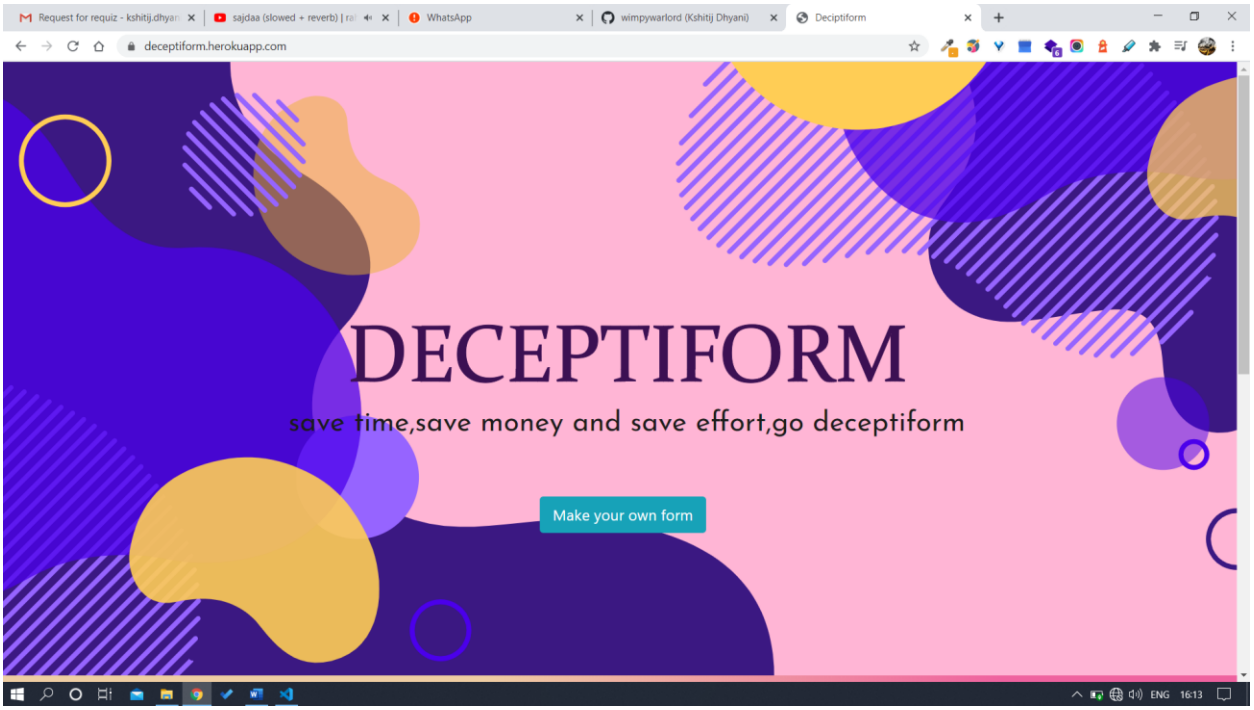
1. DECEPTIFORM

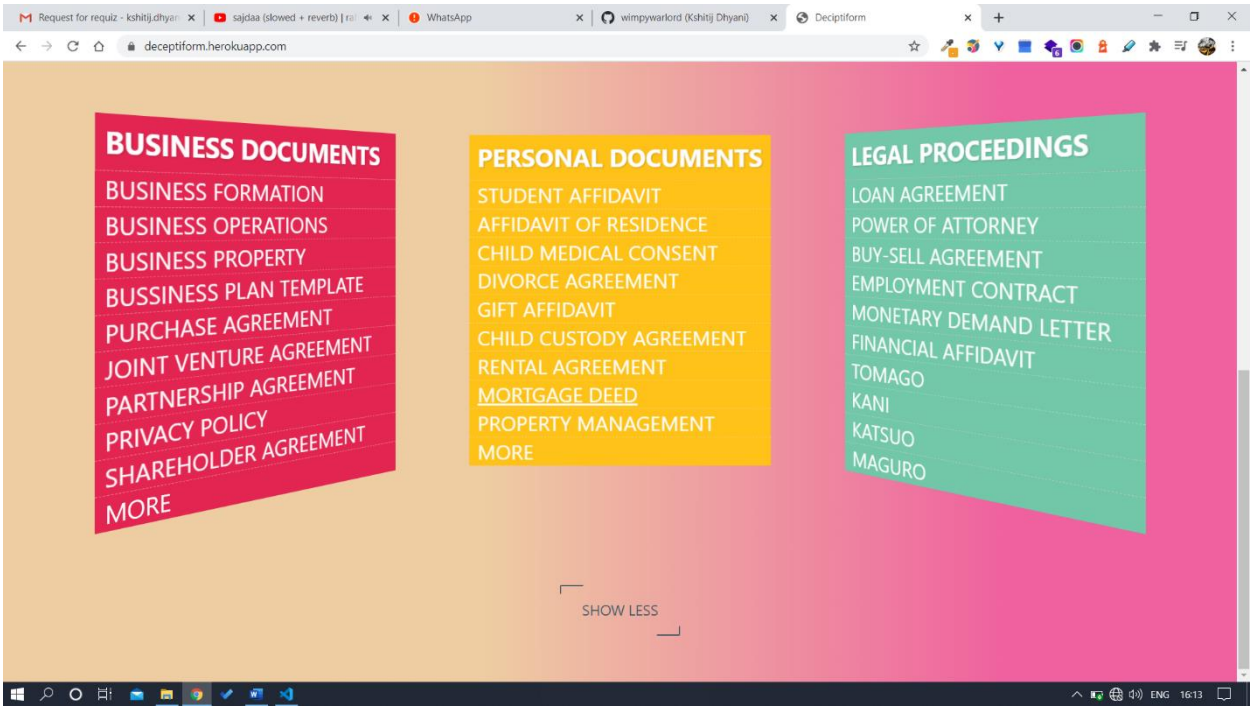
```
var express = require("express");
var bodyParser = require("body-parser");
var fs=require('fs');
var deepai=require('deepai');

deepai.setApiKey('quickstart-QUdJIGlzIGNvbWluZy4uLi4K');
var resp;
async function asyncall() {
    resp = await deepai.callStandardApi("summarization", {
            text: fs.createReadStream("proxies.txt"),
    });
    console.log(resp);

}

asyncall();
var global_variable_for_the_form_text;

var app = express();
app.use(bodyParser.urlencoded({extended : true}));
app.use(express.static(__dirname + '/assests'));
app.use(express.static(__dirname + '/public'));









////

var express = require("express");
var bodyParser = require("body-parser");




var global_variable_for_the_form_text;

var app = express();
```

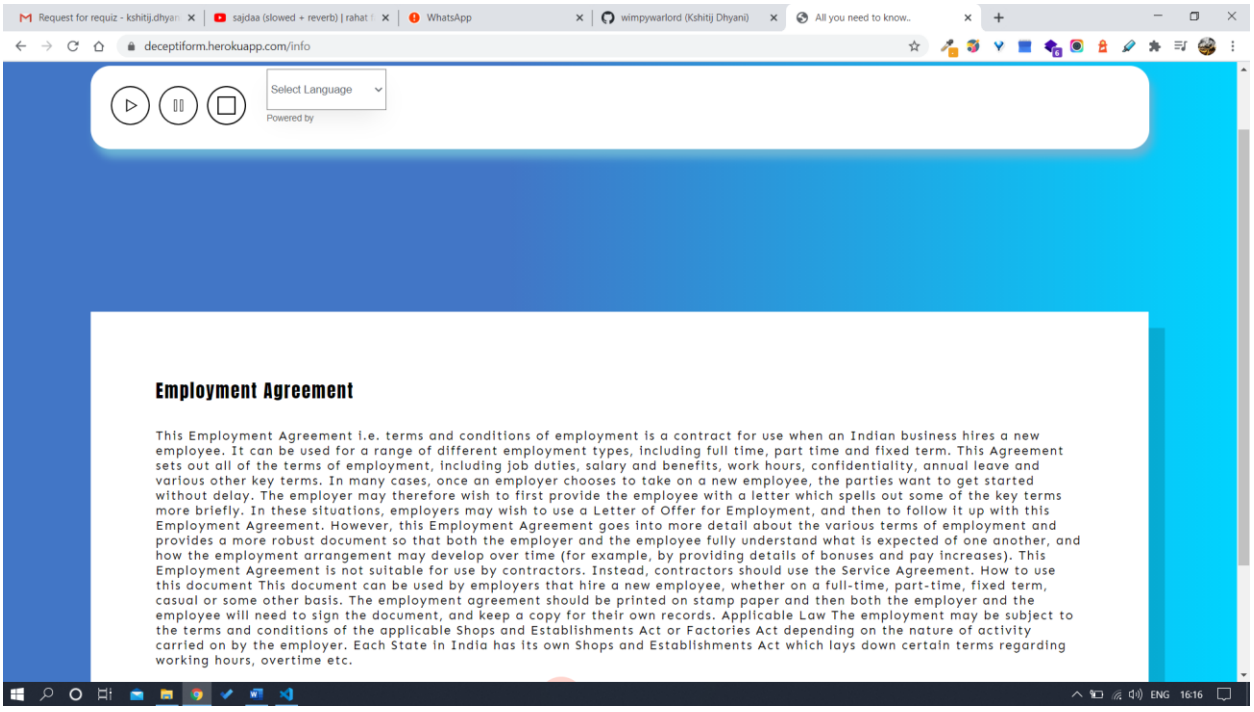
```
app.use(bodyParser.urlencoded({extended : true}));
app.use(express.static(__dirname + '/assests'));
app.use(express.static(__dirname + '/public'));

var unirest = require("unirest");

var req = unirest("POST", "https://textanalysis-keyword-extraction-
v1.p.rapidapi.com/keyword-extractor-text");

req.headers({
    "x-rapidapi-host": "textanalysis-keyword-extraction-v1.p.rapidapi.com",
    "x-rapidapi-key": "7703ef81e0mshd5555ce8567e172p148f04jsn2f562e2f9571",
    "content-type": "application/x-www-form-urlencoded"
});

req.form({
    "text": "____ (NAME)"
    +"____"

    +"____ (DATE)"

    +"____ (EMAIL)"

    +"To _____(COMPANY),"
    +"I am formally requesting you to give me _____(NO OF DAYS) days leave of abs
ence."
    +"I have finished almost all the projects assigned to me with only one remain
ing which I have given to my assistant to take care of it since he is as experien
ced just as I am in handling that remaining project. "


    +"Thank you for your consideration."

    +"Yours Sincerely,"
    +"_____(NAME)",
    "wordnum": "5"
});
var a,b;
req.end(function (res) {
    if (res.error) throw new Error(res.error);
    var keywords_arr = ["name","date of leave","company name","email"];
    res.body.keywords =  (res.body.keywords).concat(keywords_arr);
    console.log(res.body);
    var questionable = require('questionable');
questionable("employee name", function (err, titles) {
```

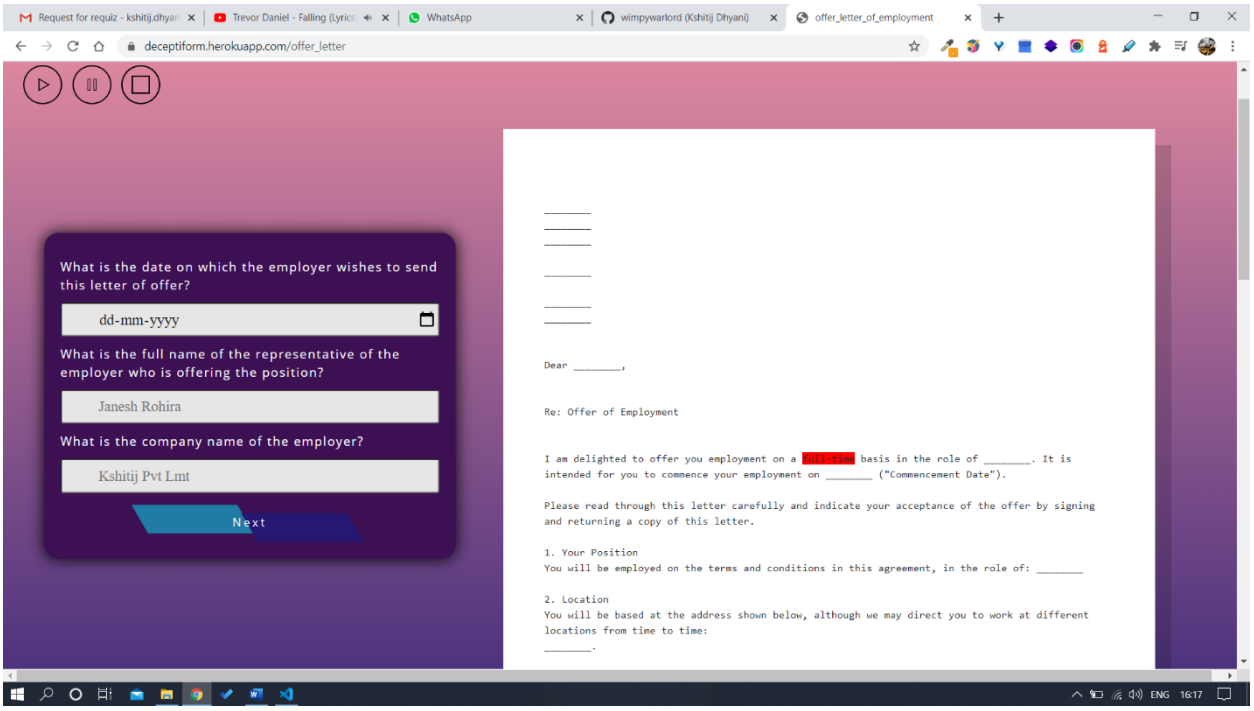
```
  if (err) throw err;
  //console.log(titles);
  console.log(titles)

    });
questionable("company name", function (err, titles) {
  if (err) throw err;
  //console.log(titles);

  console.log(titles);

    });
    questionable("date of leave", function (err, titles) {
        if (err) throw err;
        //console.log(titles);

        console.log(titles);

          });
          questionable("email", function (err, titles) {
            if (err) throw err;
            //console.log(titles);

            console.log(titles);

              });
});

app.get("/",function(req,res){
    res.render("home.ejs");
})


var questions={
    "q" : [
    "what is the name of your name?",
    "what is the name of your company?",
    "what is your name?",
    "what is the date of application of leave?"
]
}


app.get("/intake",function(req,res){
    res.render("intake.ejs");
```

```
})

app.get("/info",function(req,res){
    res.render("form_des.ejs");
})

app.get("/offer_letter",function(req,res){
    res.render("offer_letter.ejs",{summary:resp.output});
})

app.get("/autoform",function(req,res){
    res.render("autoform.ejs",{questions:questions});
})

app.listen( process.env.PORT || 8000 , function(){
    console.log("SERVER 8000 HAS STARTED");
});
```

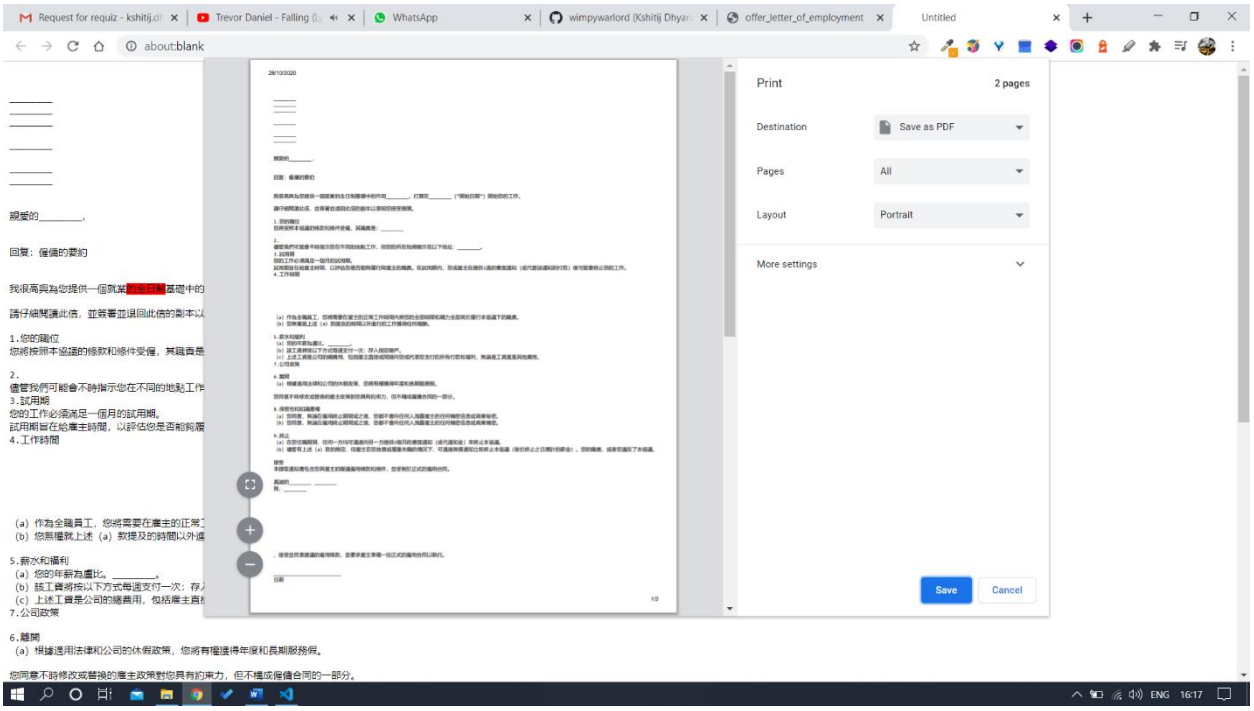



deceptiform.herokuapp.com
BUSINESS DOCUMENTS
BUSINESS FORMATION
BUSINESS OPERATIONS
BUSINESS PROPERTY
BUSSINESS PLAN TEMPLATE
PURCHASE AGREEMENT
JOINT VENTURE AGREEMENT
PARTNERSHIP AGREEMENT
PRIVACY POLICY
SHAREHOLDER AGREEMENT
MORE
PERSONAL DOCUMENTS
STUDENT AFFIDAVIT
AFFIDAVIT OF RESIDENCE
CHILD MEDICAL CONSENT
DIVORCE AGREEMENT
GIFT AFFIDAVIT
CHILD CUSTODY AGREEMENT
RENTAL AGREEMENT
MORTGAGE DEED
PROPERTY MANAGEMENT
MORE
LEGAL PROCEEDINGS
LOAN AGREEMENT
POWER OF ATTORNEY
BUY-SELL AGREEMENT
EMPLOYMENT CONTRACT
MONETARY DEMAND LETTER
FINANCIAL AFFIDAVIT
TOMAGO
KANI
KATSUO
MAGURO
SHOW LESS


deceptiform.herokuapp.com/info
Select Language
Powered by
Employment Agreement
This Employment Agreement i.e. terms and conditions of employment is a contract for use when an Indian business hires a new employee. It can be used for a range of different employment types, including full time, part time and fixed term. This Agreement sets out all of the terms of employment, including job duties, salary and benefits, work hours, confidentiality, annual leave and various other key terms. In many cases, once an employer chooses to take on a new employee, the parties want to get started without delay. The employer may therefore wish to first provide the employee with a letter which spells out some of the key terms more briefly. In these situations, employers may wish to use a Letter of Offer for Employment, and then to follow it up with this Employment Agreement. However, this Employment Agreement goes into more detail about the various terms of employment and provides a more robust document so that both the employer and the employee fully understand what is expected of one another, and how the employment arrangement may develop over time (for example, by providing details of bonuses and pay increases). This Employment Agreement is not suitable for use by contractors. Instead, contractors should use the Service Agreement. How to use this document This document can be used by employers that hire a new employee, whether on a full-time, part-time, fixed term, casual or some other basis. The employment agreement should be printed on stamp paper and then both the employer and the employee will need to sign the document, and keep a copy for their own records. Applicable Law The employment may be subject to the terms and conditions of the applicable Shops and Establishments Act or Factories Act depending on the nature of activity carried on by the employer. Each State in India has its own Shops and Establishments Act which lays down certain terms regarding working hours, overtime etc.


Request for requiz - kshitij.dhyan
Trevor Daniel - Falling (Lyrics
WhatsApp
wimpywarlord (Kshitij Dhyani)
offer_letter_of_employment
deceptiform.herokuapp.com/offer_letter
What is the date on which the employer wishes to send this letter of offer?
dd-mm-yyyy
What is the full name of the representative of the employer who is offering the position?
Janesh Rohira
What is the company name of the employer?
Kshitij Pvt Lmt
Next
_____
_____
_____
_____
_____
_____
Dear _____,
Re: Offer of Employment
I am delighted to offer you employment on a full-time basis in the role of _____. It is intended for you to commence your employment on _____ ("Commencement Date").
Please read through this letter carefully and indicate your acceptance of the offer by signing and returning a copy of this letter.
1. Your Position
You will be employed on the terms and conditions in this agreement, in the role of: _____
2. Location
You will be based at the address shown below, although we may direct you to work at different locations from time to time:
_____.
ENG
16:17


Request for requiz - kshitij.dhyan
Trevor Daniel - Falling (Lyrics
WhatsApp
wimpywarlord (Kshitij Dhyani)
offer_letter_of_employment
deceptiform.herokuapp.com/offer_letter
Google
Translated to: Chinese (Traditional)
Show original
Options
雇主希望在什麼日期發送此錄取通知書?
dd-mm-yyyy
提供職位的雇主代表的全名是什麼?
珍妮絲·羅希拉 (Janesh Rohira)
雇主的公司名稱是什麼?
Kshitij Pvt Lmt
下一個
請仔細閱讀此信，並簽署並退回此信的副本以表明您接受報價。
1.您的職位
您將按照本協議的條款和條件受僱，其職責是：_____
2.
儘管我們可能會不時指示您在不同的地點工作，但您的所在地將顯示在以下地址：_____。
3.試用期
您的工作必須滿足一個月的試用期。
試用期旨在給雇主時間，以評估您是否能夠履行與雇主的職責。在試用期內，您或雇主在提供1週的書面通知（或代替該通知的付款）後可能會終止您的工作。
4.工作時間
(a) 作為全職員工，您將需要在雇主的正常工作時間內將您的全部時間和精力全部用於履行本協議下的職責。
(b) 您無權就上述 (a) 款提及的時間以外進行的工作獲得任何報酬。
5.薪水和福利
(a) 您的年薪為盧比。_____。
(b) 該工資將按以下方式每週支付一次：存入指定賬戶。
(c) 上述工資是公司的總費用，包括雇主直接或間接向您或代表您支付的所有付款和福利，無論是工資還是其他費用。
7.公司政策
6.離開
(a) 根據適用法律和公司的休假政策，您將有權獲得年度和長期服務假。
您同意不時修改或替換的雇主政策對您具有約束力，但不構成僱傭合同的一部分。
8.保密性和知識產權
(a) 您同意，無論在僱用終止期間或之後，您都不會向任何人洩露雇主的任何機密信息或商業秘密。
(b) 您同意，無論在僱用終止期間或之後，您都不會向任何人洩露雇主的任何機密信息或商業機密。
ENG
16:17

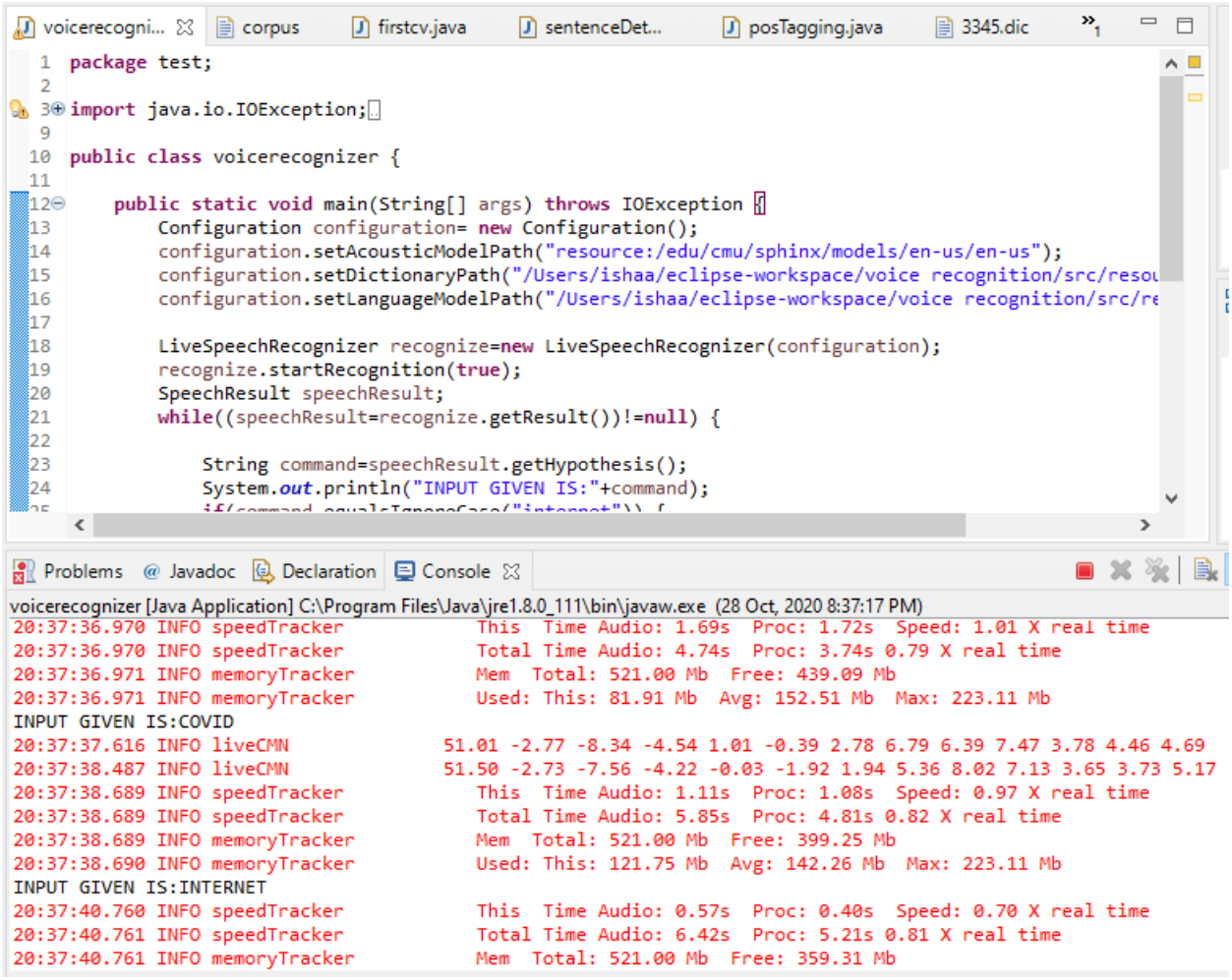

Request for requiz - kshitij.dh
Trevor Daniel - Falling
WhatsApp
wimpywarlord (Kshitij Dhyani
offer_letter_of_employment
Untitled
about:blank
Print
2 pages
Destination
Save as PDF
Pages
All
Layout
Portrait
More settings
Save
Cancel
ENG
16:17

2. SPEECHR

```
package test;

import java.io.IOException;
import java.awt.Desktop;
import java.net.URI;
import edu.cmu.sphinx.api.Configuration;
import edu.cmu.sphinx.api.LiveSpeechRecognizer;
import edu.cmu.sphinx.api.SpeechResult;

public class voicerecognizer {

	public static void main(String[] args) throws IOException {
		Configuration configuration= new Configuration();

	configuration.setAcousticModelPath("resource:/edu/cmu/sp
hinx/models/en-us/en-us");
		configuration.setDictionaryPath("/Users/ishaa/eclipse-
workspace/voice recognition/src/resources/3345.dic");

	configuration.setLanguageModelPath("/Users/ishaa/eclipse
-workspace/voice recognition/src/resources/3345.lm");

		LiveSpeechRecognizer recognize=new
LiveSpeechRecognizer(configuration);
		recognize.startRecognition(true);
```

```
            SpeechResult speechResult;
            while((speechResult=recognize.getResult())!=null) {

                String command=speechResult.getHypothesis();
                System.out.println("INPUT GIVEN
IS:"+command);
                if(command.equalsIgnoreCase("internet")) {
                    try {

                        URI uri= new
URI("http://www.google.com");


java.awt.Desktop.getDesktop().browse(uri);


                    } catch (Exception e) {

                     e.printStackTrace();
                    }
                }

                }
            }
        }
```

When input is given as internet, google chrome opens up with google as the web page being opened,



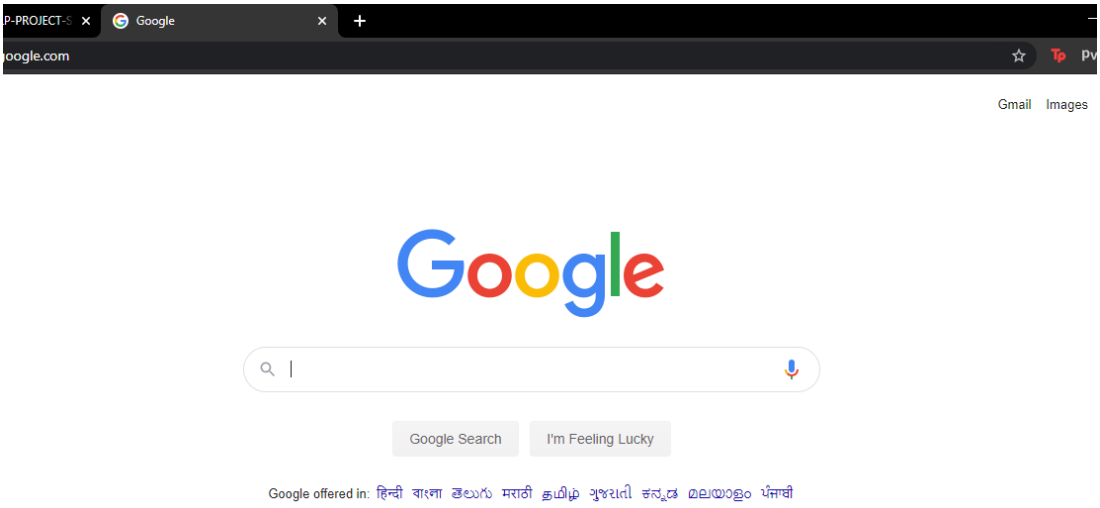
voicerecogni... | corpus | firstcv.java | sentenceDet... | posTagging.java | 3345.dic

```
1  package test;
2
3  import java.io.IOException;
9
10 public class voicerecognizer {
11
12     public static void main(String[] args) throws IOException {
13         Configuration configuration= new Configuration();
14         configuration.setAcousticModelPath("resource:/edu/cmu/sphinx/models/en-us/en-us");
15         configuration.setDictionaryPath("/Users/ishaa/eclipse-workspace/voice recognition/src/resou
16         configuration.setLanguageModelPath("/Users/ishaa/eclipse-workspace/voice recognition/src/re
17
18         LiveSpeechRecognizer recognize=new LiveSpeechRecognizer(configuration);
19         recognize.startRecognition(true);
20         SpeechResult speechResult;
21         while((speechResult=recognize.getResult())!=null) {
22
23             String command=speechResult.getHypothesis();
24             System.out.println("INPUT GIVEN IS:"+command);
```

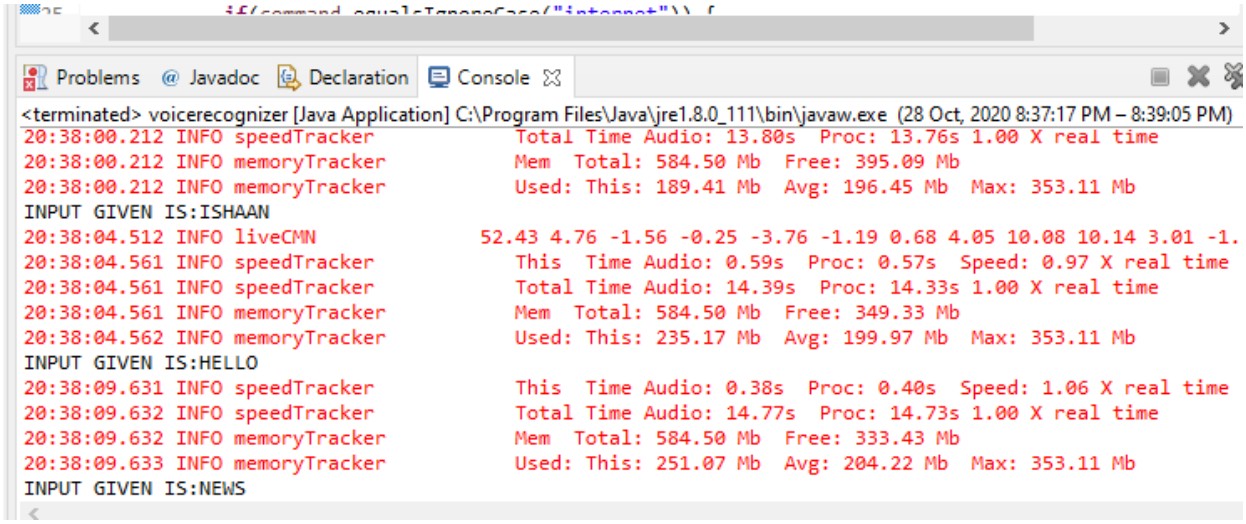
Problems | Javadoc | Declaration | Console

voicerecognizer [Java Application] C:\Program Files\Java\jre1.8.0_111\bin\javaw.exe (28 Oct, 2020 8:37:17 PM)

```
20:37:36.970 INFO speedTracker            This  Time Audio: 1.69s  Proc: 1.72s  Speed: 1.01 X real time
20:37:36.970 INFO speedTracker            Total Time Audio: 4.74s  Proc: 3.74s 0.79 X real time
20:37:36.971 INFO memoryTracker           Mem  Total: 521.00 Mb  Free: 439.09 Mb
20:37:36.971 INFO memoryTracker           Used: This: 81.91 Mb  Avg: 152.51 Mb  Max: 223.11 Mb
INPUT GIVEN IS:COVID
20:37:37.616 INFO liveCMN              51.01 -2.77 -8.34 -4.54 1.01 -0.39 2.78 6.79 6.39 7.47 3.78 4.46 4.69
20:37:38.487 INFO liveCMN              51.50 -2.73 -7.56 -4.22 -0.03 -1.92 1.94 5.36 8.02 7.13 3.65 3.73 5.17
20:37:38.689 INFO speedTracker            This  Time Audio: 1.11s  Proc: 1.08s  Speed: 0.97 X real time
20:37:38.689 INFO speedTracker            Total Time Audio: 5.85s  Proc: 4.81s 0.82 X real time
20:37:38.689 INFO memoryTracker           Mem  Total: 521.00 Mb  Free: 399.25 Mb
20:37:38.690 INFO memoryTracker           Used: This: 121.75 Mb  Avg: 142.26 Mb  Max: 223.11 Mb
INPUT GIVEN IS:INTERNET
20:37:40.760 INFO speedTracker            This  Time Audio: 0.57s  Proc: 0.40s  Speed: 0.70 X real time
20:37:40.761 INFO speedTracker            Total Time Audio: 6.42s  Proc: 5.21s 0.81 X real time
20:37:40.761 INFO memoryTracker           Mem  Total: 521.00 Mb  Free: 359.31 Mb
```



Speech being recognized,



Corpus,

```
open
internet
school
hello
ishaan
covid
news
```

CONCLUSION

This future of Deceptiform aims to make a central database of all the forms that are generally used in the country and make it a cakewalk to fill such forms. Giving a time benefit of a magnitude much greater than that of current standards. Save time, save money, and save the effort, go deceptiform.

REFERENCES

Deceptiform website: https://deceptiform.herokuapp.com/
Github repository: https://github.com/wimpywarlord/women_techies
Text to speech api : https://cloud.google.com/text-to-speech
Text summarization api : https://rapidapi.com/textanalysis/api/text-summarization