CS 4624 Multimedia, Hypertext, and Information Access - Dr. Fox
Virginia Tech, Blacksburg, VA 24061
5/10/22

Risk Prediction Sentiment Analysis 3

Kaustubh Kale, Jaswant Kasinedi, Sagar Mehta, Parvesh Samayamanthula, Arjun Vashistha

Outline

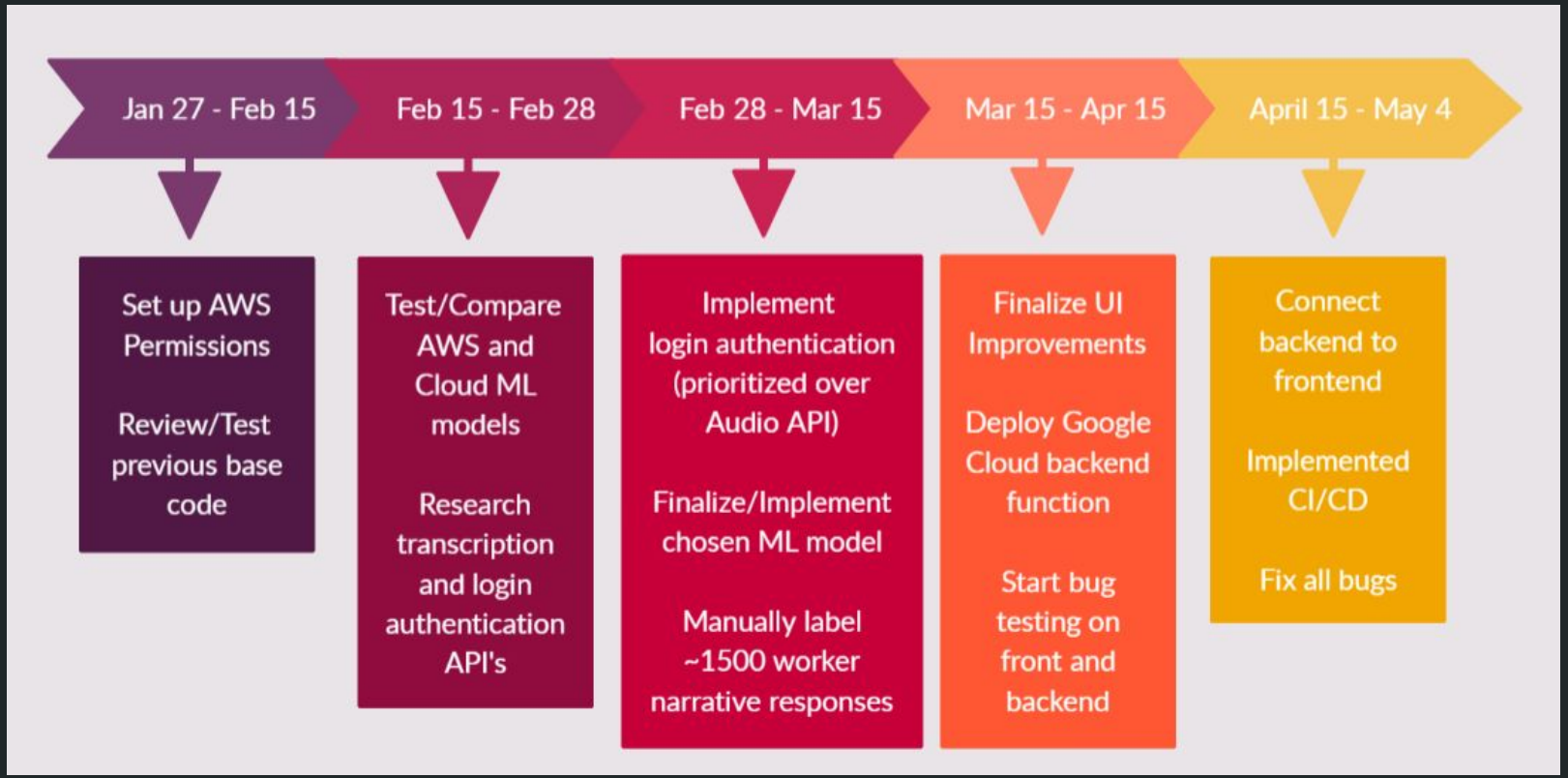
- Project Overview
- Timeline
- Frontend
- Backend
- Lessons Learned
- Future Work
- Acknowledgments
- References

Project Overview

- Hypothesis: Worker sentiment in the workplace indicates the overall safety of the workplace
- Intuitive web app for users to upload worker responses and get sentiment analysis results

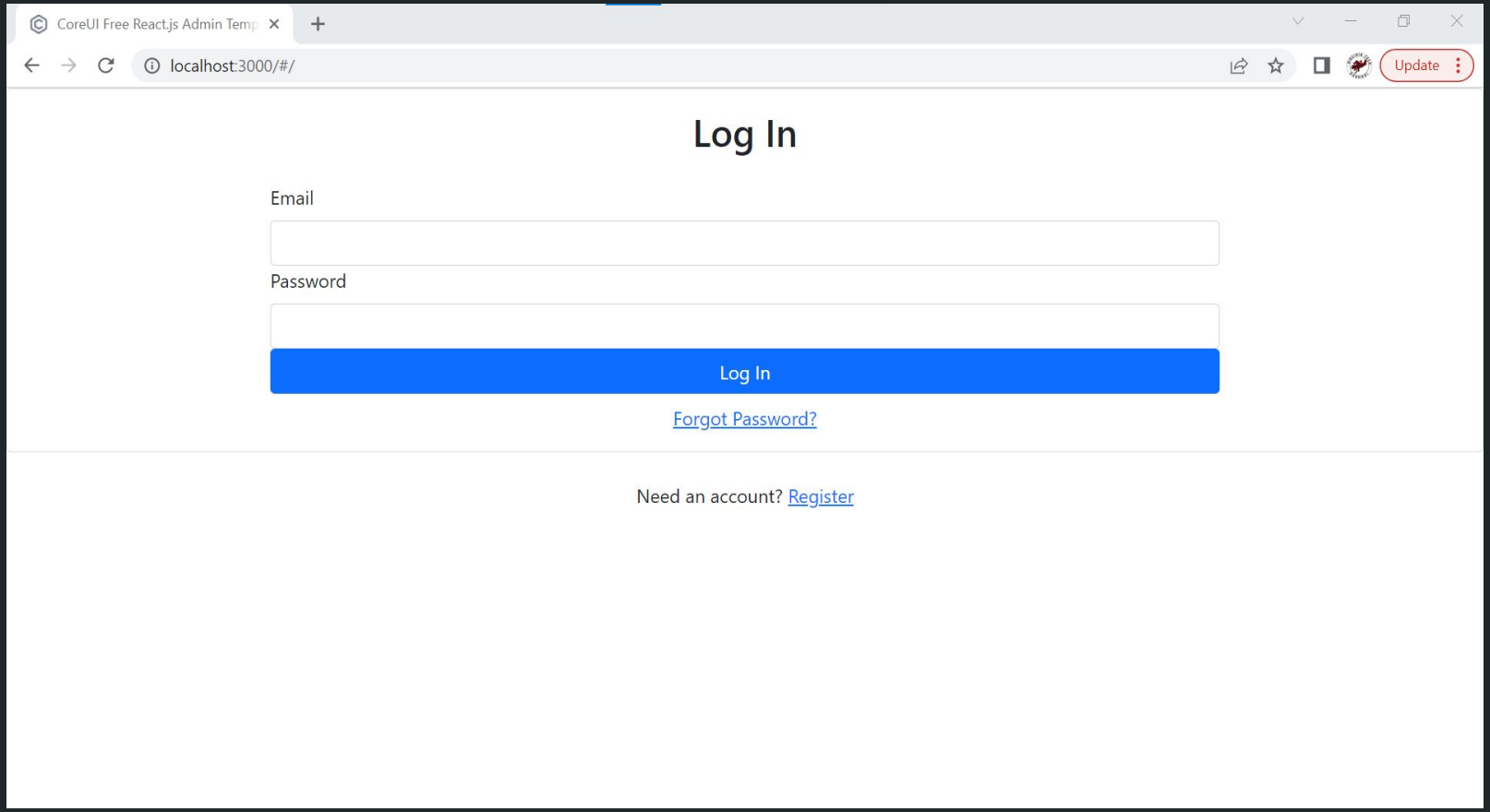
Client

Timeline

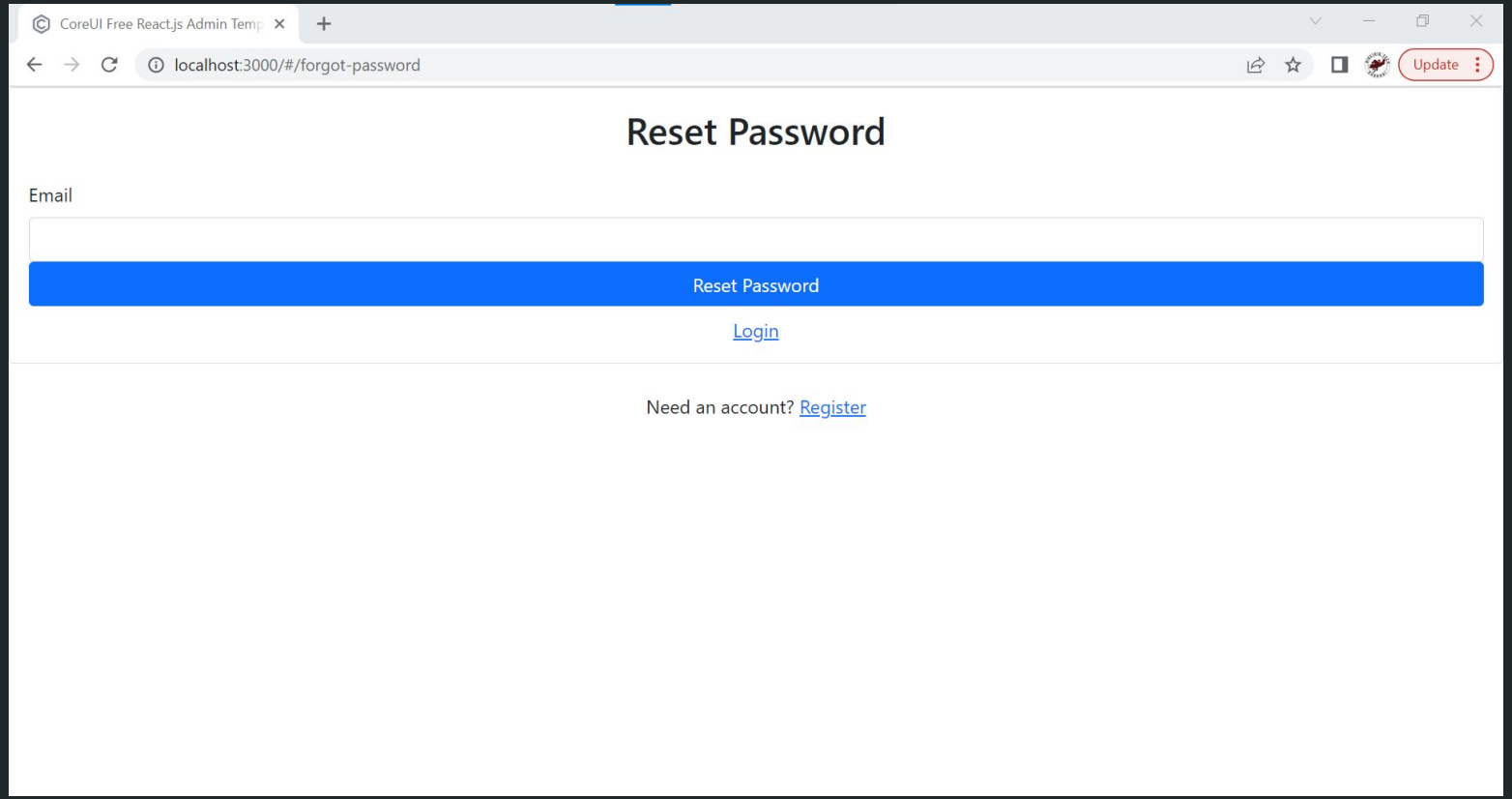

Jan 27 - Feb 15
Feb 15 - Feb 28
Feb 28 - Mar 15
Mar 15 - Apr 15
April 15 - May 4
Set up AWS Permissions
Review/Test previous base code
Test/Compare AWS and Cloud ML models
Research transcription and login authentication API's
Implement login authentication (prioritized over Audio API)
Finalize/Implement chosen ML model
Manually label ~1500 worker narrative responses
Finalize UI Improvements
Deploy Google Cloud backend function
Start bug testing on front and backend
Connect backend to frontend
Implemented CI/CD
Fix all bugs

Frontend

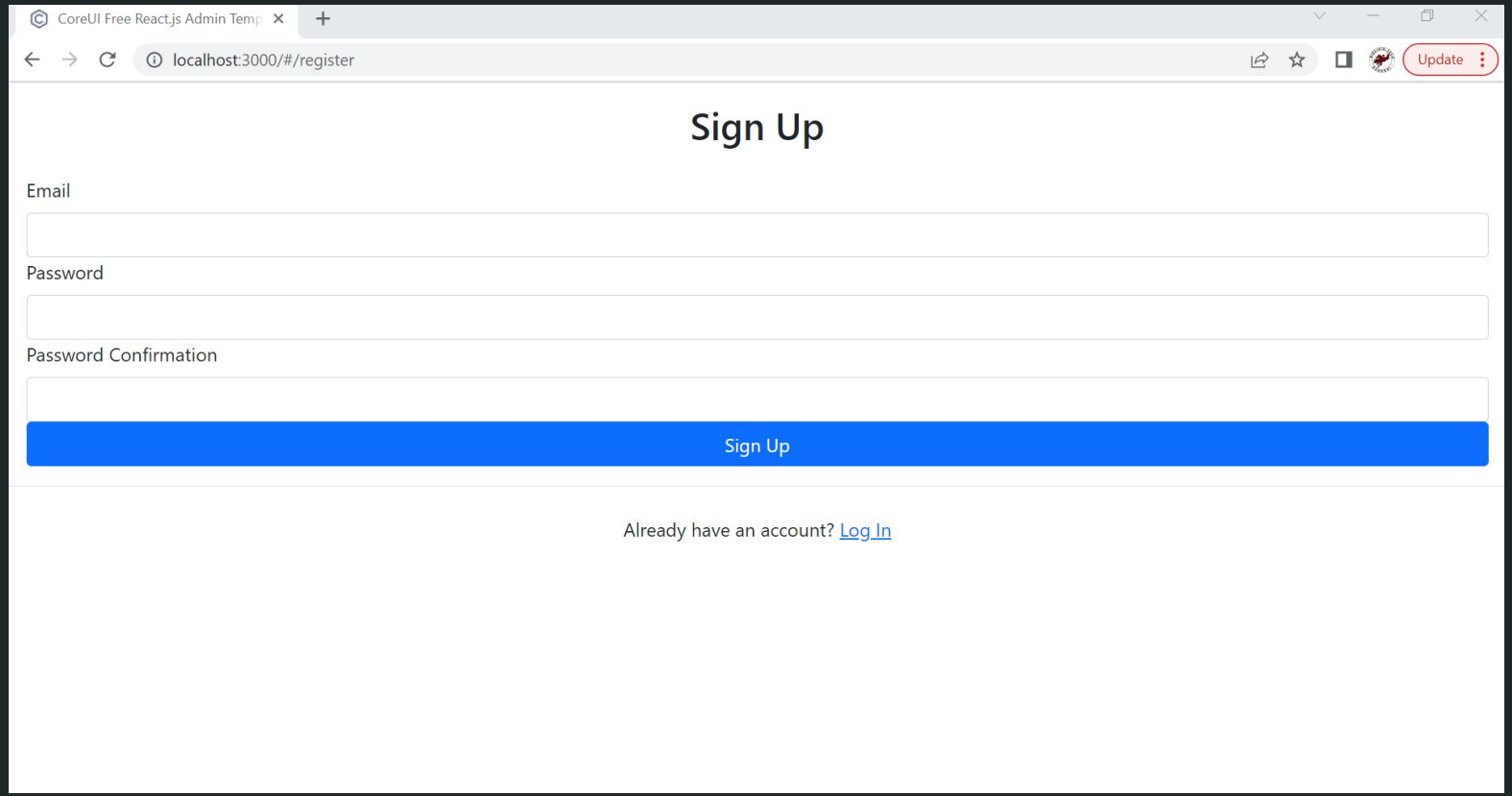
Log In


CoreUI Free React.js Admin Temp
localhost:3000/#/
Update
Log In
Email
Password
Log In
Forgot Password?
Need an account? Register

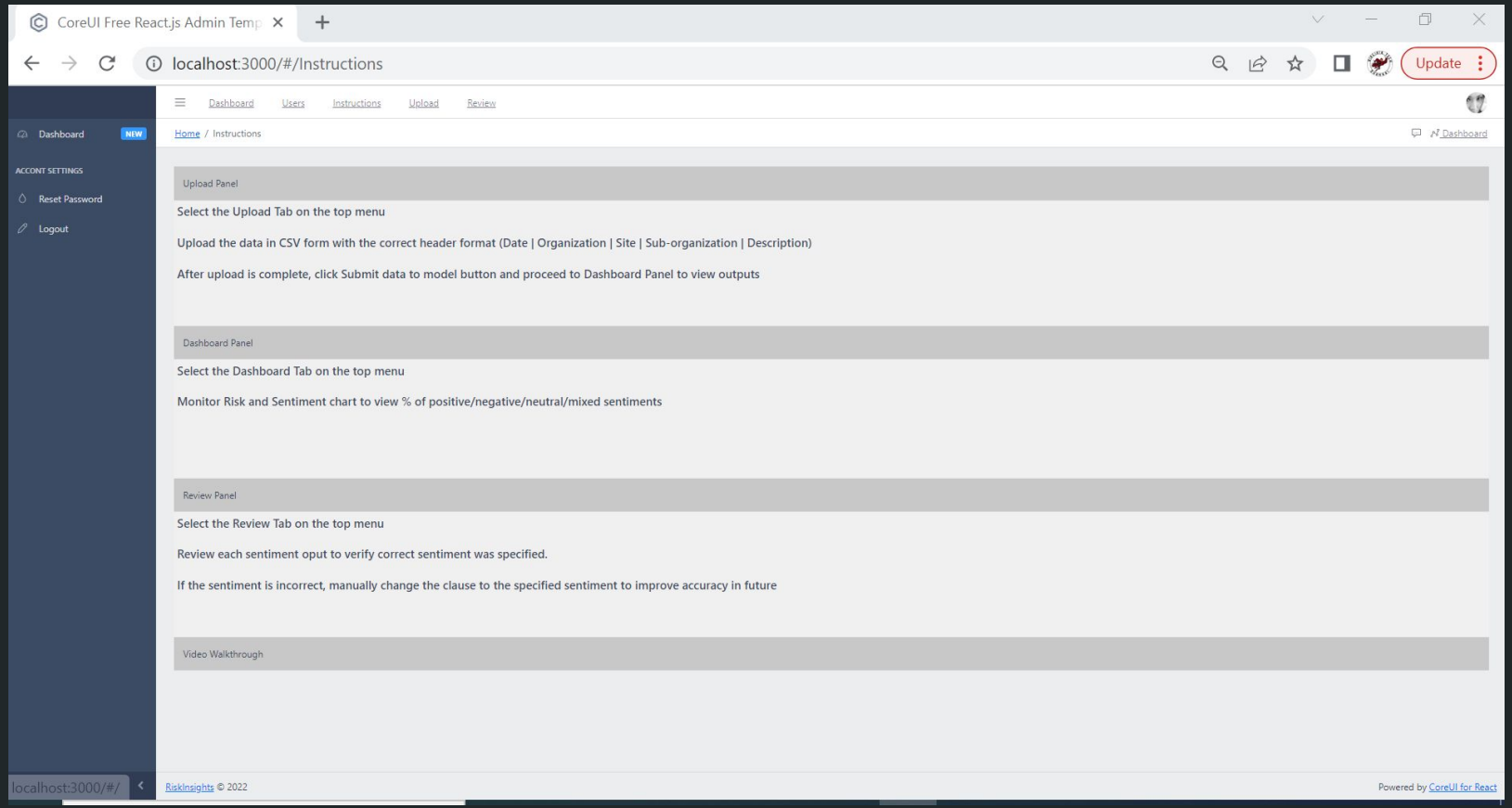
Reset Password



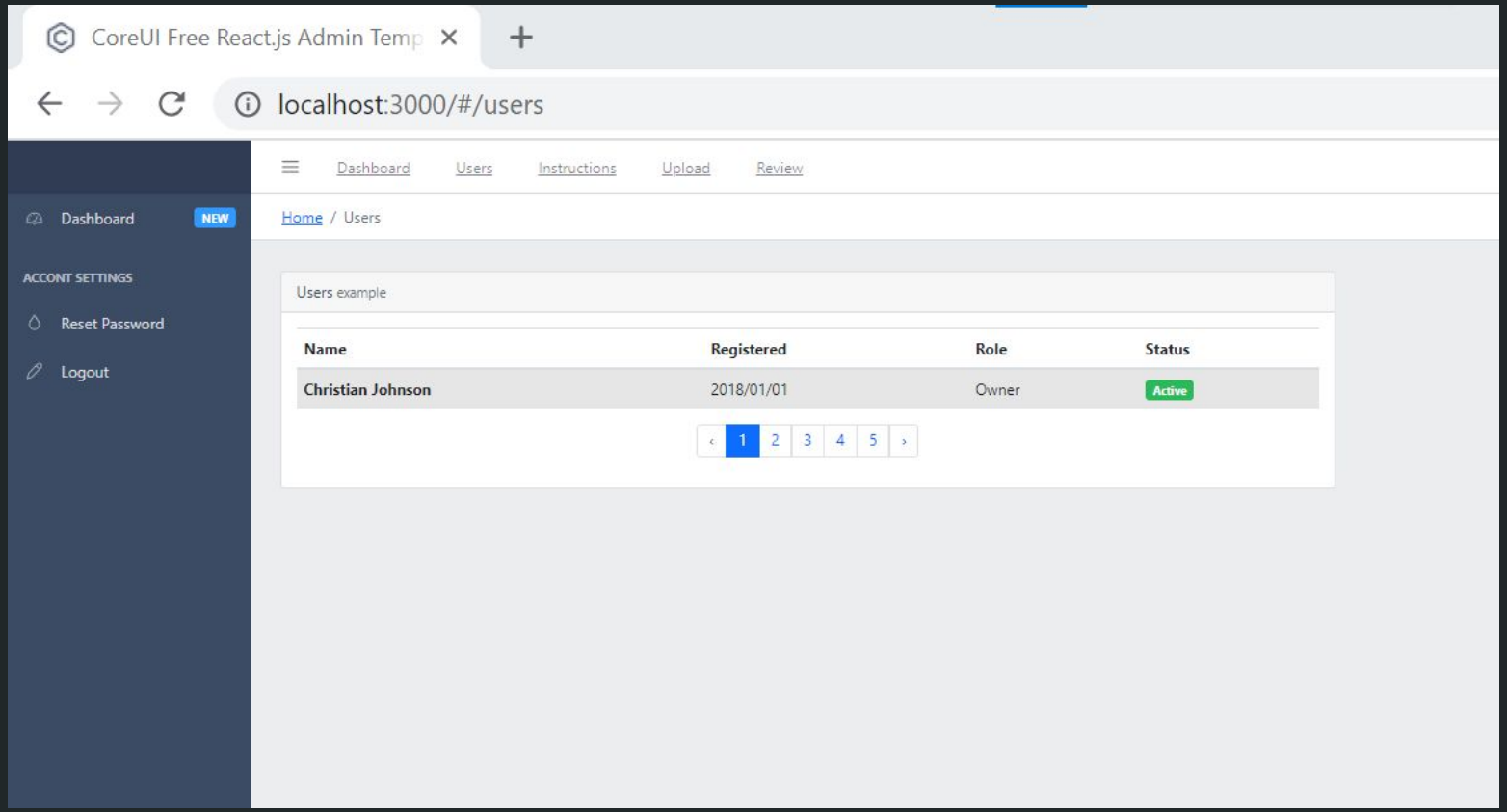
Sign Up


CoreUI Free React.js Admin Temp
localhost:3000/#/register
Update
Sign Up
Email
Password
Password Confirmation
Sign Up
Already have an account? Log In

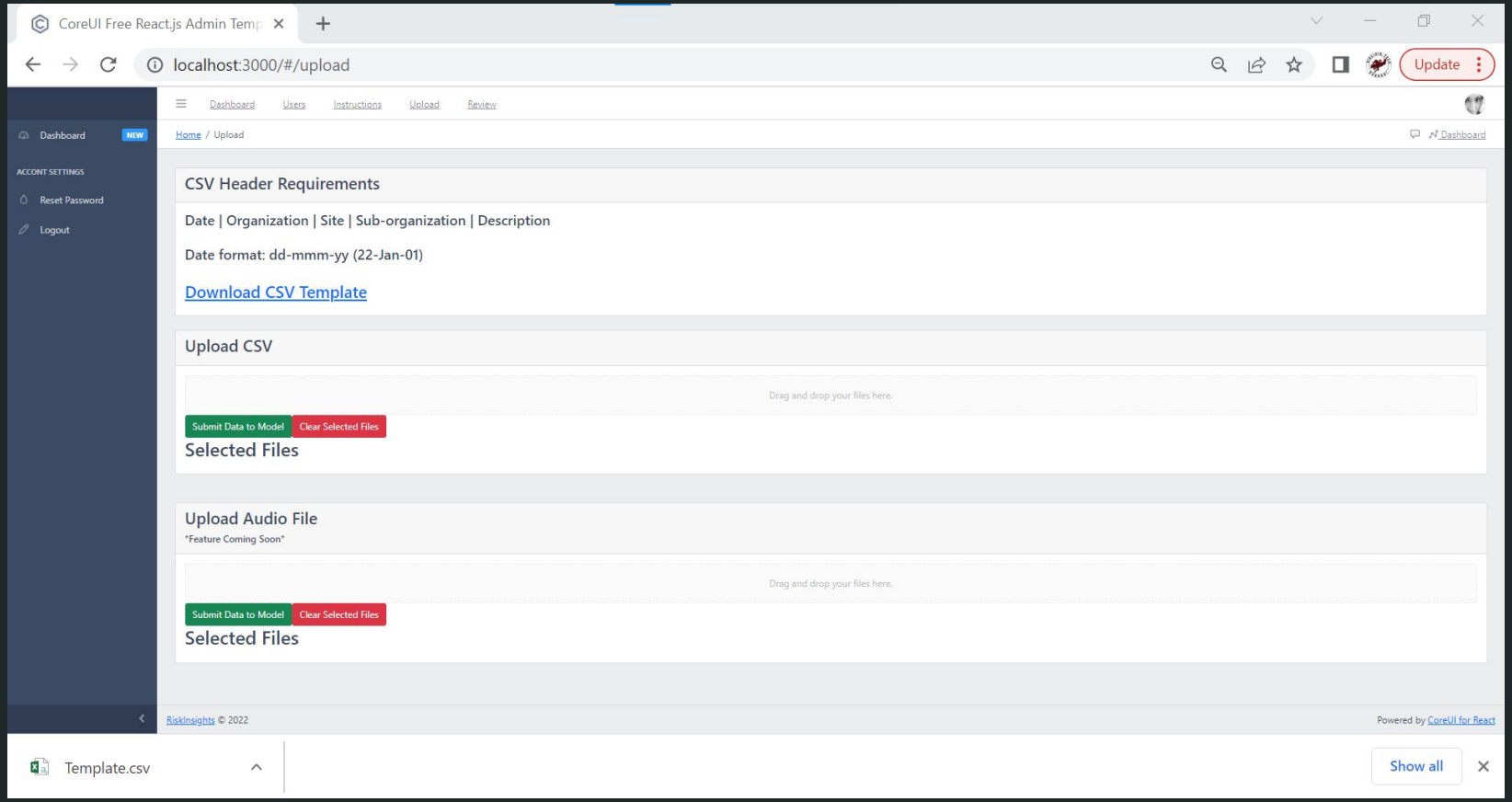
Instructions



Users


CoreUI Free React.js Admin Temp
localhost:3000/#/users
Dashboard
Users
Instructions
Upload
Review
Dashboard
NEW
ACCONT SETTINGS
Reset Password
Logout
Home / Users
Users example
Name
Registered
Role
Status
Christian Johnson
2018/01/01
Owner
Active
1
2
3
4
5

Upload

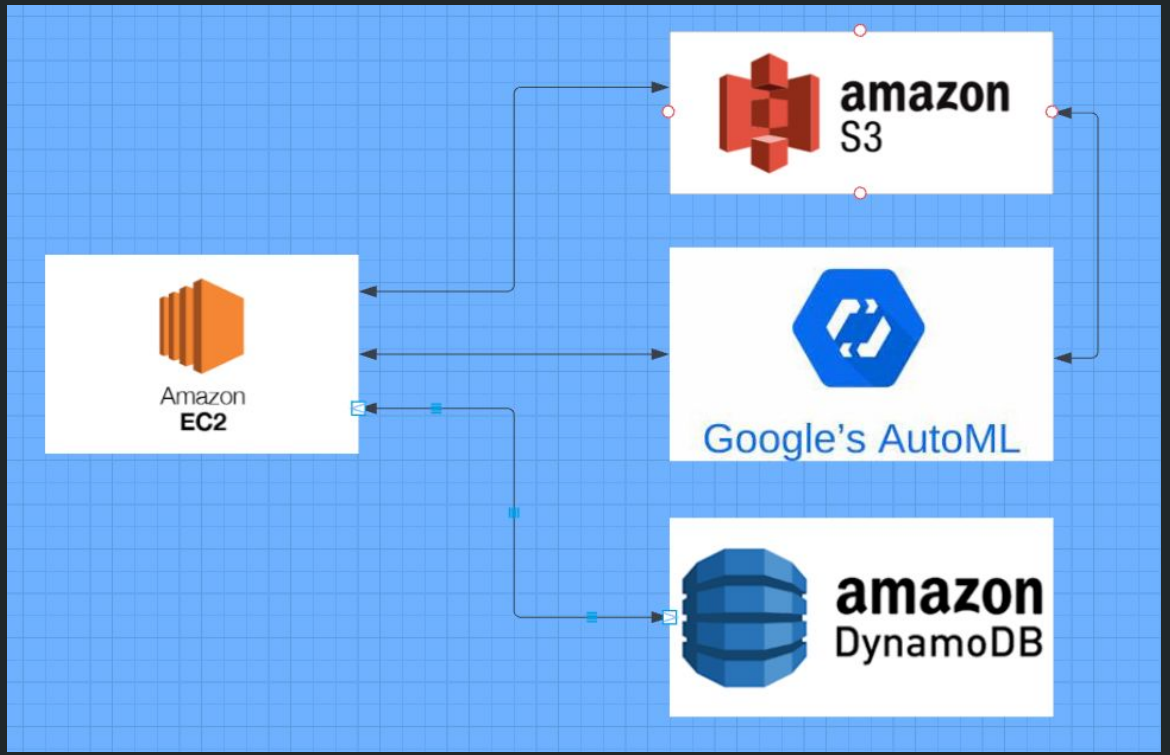

CoreUI Free React.js Admin Temp
localhost:3000/#/upload
Update
Dashboard
Users
Instructions
Upload
Review
Dashboard NEW
ACCONT SETTINGS
Reset Password
Logout
Home / Upload
Dashboard
CSV Header Requirements
Date | Organization | Site | Sub-organization | Description
Date format: dd-mmm-yy (22-Jan-01)
Download CSV Template
Upload CSV
Drag and drop your files here.
Submit Data to Model
Clear Selected Files
Selected Files
Upload Audio File
Feature Coming Soon
Drag and drop your files here.
Submit Data to Model
Clear Selected Files
Selected Files
RiskInsights © 2022
Powered by CoreUI for React
Template.csv
Show all

Backend

Sentiment Analysis

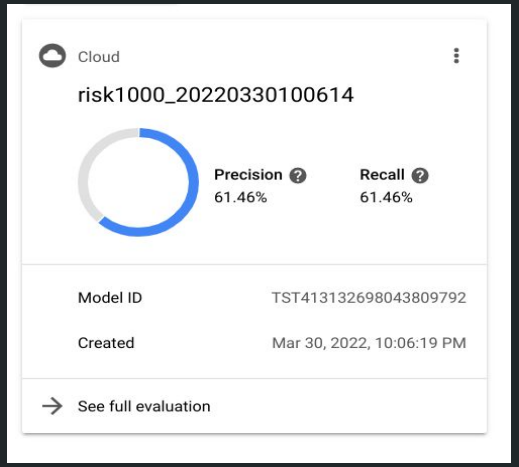
- Services tested: Microsoft Azure, Google Cloud Platform
- Chose GCP for its custom AutoML Service
- Accepts range of sentiment scores and labeled data (Ex: 0-4)

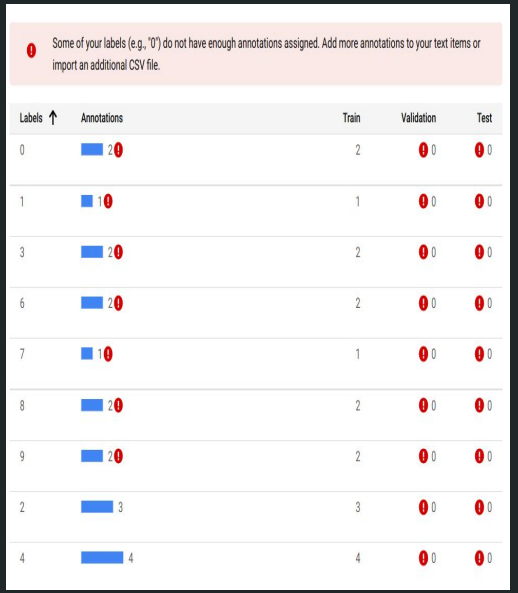
Backend Infrastructure



AutoML Sentiment Analysis

- Labeled 1544 worker responses for supervised learning
- Initially tried a sentiment range of 0-10
- Settled on a range of 0-4





Revised Accuracy Evaluation

- Accuracy of predicting whether worker responses are positive or negative

```
accurate = 0
[accurate := accurate + 1 for m, r in zip(model_ratings, test_ratings) if m == 2 and r == 2 or (
    (m == 3 or m == 4) and (r == 3 or r == 4)) or ((m == 0 or m == 1) and (r == 0 or r == 1))]

print(accurate / len(test_ratings))
```

```
0.8469387755102041
```

Input.csv

Date	Organization	Site	Sub-organization	Description
2018-09-18	Global Operations	Site B	Wetmill	Asked Wetmill/Feedhouse team if machine guarding audits have been completed. Not aware of a recent audit. Removable gate in question should have a lock on it. Notified Area Technologist and Area Engineer and asked that the chain be replaced at the DSL dryer.
2018-10-04	Global Operations	Site B	Office area	Talked to employee about their immediate work space and the ergonomic impact of monitor position.
2018-10-20	Global Operations	Site C	Operations	Great job and clean up

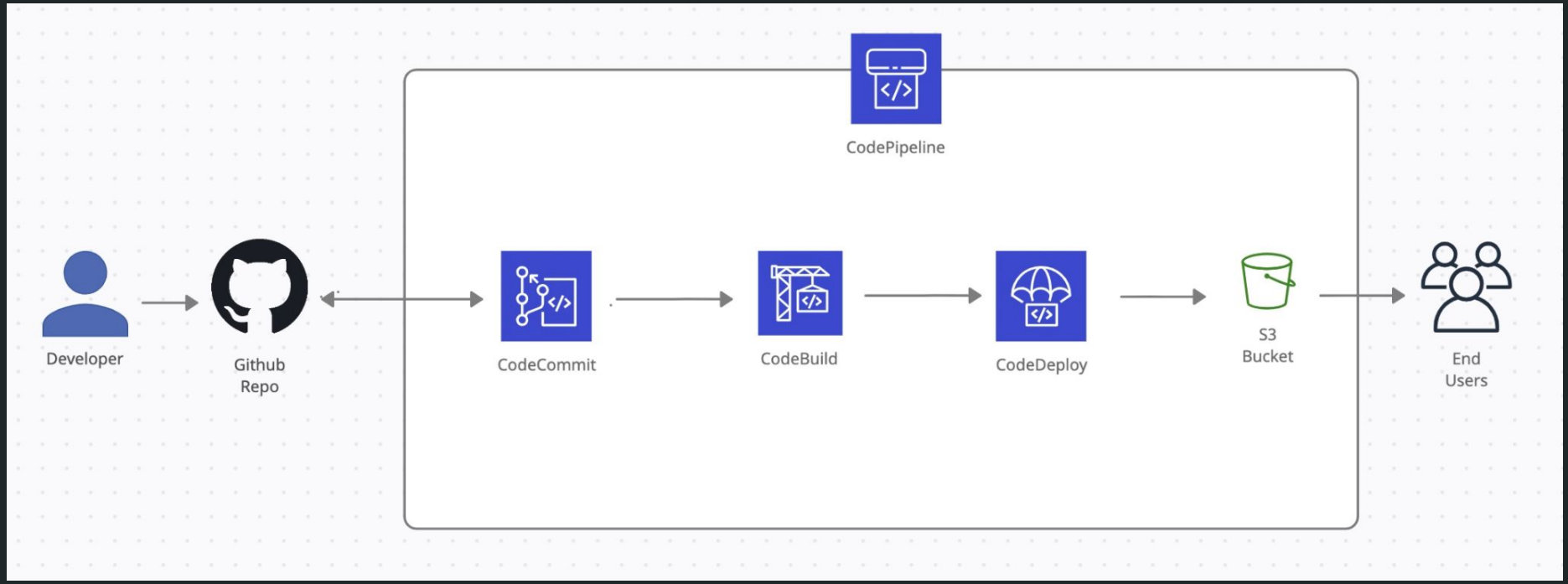
Upload Page Workflow

- User uploads input CSV to upload page
- CSV gets sent to the backend, is parsed, and worker descriptions are sent to the Google Cloud model
- Sentiment results are sent back from the model and appended to input CSV
- New output CSV is sent back to frontend, where it’s automatically downloaded

Output.csv

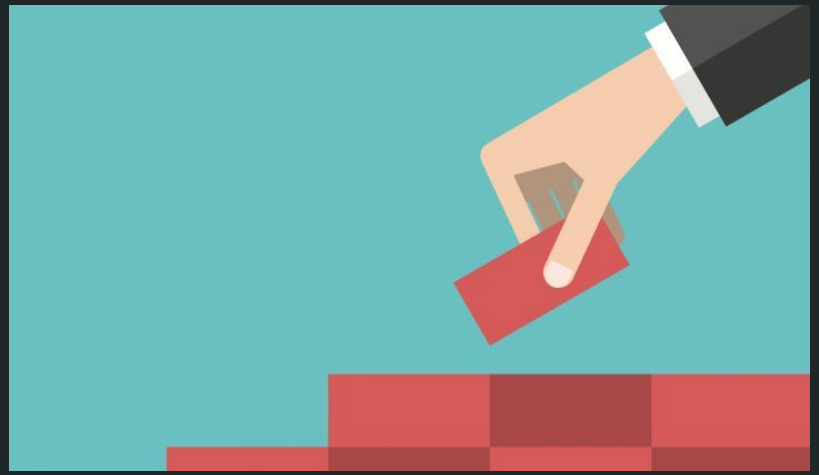
Date	Organization	Site	Sub-organization	Description	Sentiment Score	Overall Sentiment
2018-09-18	Global Operations	Site B	Wetmill	Asked Wetmill/Feedhouse team if machine guarding audits have been completed. Not aware of a recent audit. Removable gate in question should have a lock on it. Notified Area Technologist and Area Engineer and asked that the chain be replaced at the DSL dryer.	2	Neutral
2018-10-04	Global Operations	Site B	Office area	Talked to employee about their immediate work space and the ergonomic impact of monitor position.	2	Neutral
2018-10-20	Global Operations	Site C	Operations	Great job and clean up	4	Positive

CI/CD


CodePipeline
Developer
Github
Repo
CodeCommit
CodeBuild
CodeDeploy
S3
Bucket
End
Users

Lessons Learned

- Time management - gauging which features to prioritize and the time to implement them
- Quality of data - skewed data results
- Building on previous team’s work and ideas



Future Work

- Audio transcription
- Continuous learning/improvement of ML model
- Testing and feedback tool
- Better visualizations UI for sentiment data

Acknowledgements

- Dr. Fox
  - Mentor
- Christian Johnson
  - Client

References

[1] 360training. New Survey Reveals Rise In Workplace Safety Concerns. https://www.360training.com/blog/new-survey-reveals-rise-workplace-safety-concerns#:~:text=Top%20Workplace%20Safety%20Concerns%20in%202019&text=Electrical%20Hazards%20(13%25),Falling%20Objects%20(6%25), 2022. Accessed: 30-Jan-2022.
[2] Oxford Language. Oxford Languages and Google. https://languages.oup.com/google-dictionary-en/, 2022.
[3] Google Firebase. Firebase Documentation. https://firebase.google.com/docs, 2022. Accessed: 10-Mar-2022.
[4] Amazon Web Services. Set Up a CI/CD Pipeline on AWS. https://aws.amazon.com/getting-started/hands-on/set-up-ci-cd-pipeline/, 2022. Accessed: 17-Apr-2022.
[5] Google Cloud. AutoML beginner's guide. https://cloud.google.com/vertex-ai/docs/beginner/beginners-guide, 2022. Accessed: 20-Feb-2022.
[6] Google Cloud. Sentiment Analysis Tutorial https://cloud.google.com/natural-language/docs/sentiment-tutorial, 2022. Accessed: 20-Feb-2022.
[7] Amazon Web Services. Bucket policy examples. https://docs.aws.amazon.com/AmazonS3/latest/userguide/example-bucket-policies.html, 2022, Accessed: 21-Feb-2022.
[8] Google Cloud. APIs and references. https://cloud.google.com/speech-to-text/docs/apis, 2022. Accessed: 21-Feb-2022.
[9] Google Firebase. Authenticate Using Google with JavaScript. https://firebase.google.com/docs/auth/web/google-signin, 2022. Accessed: 10-Mar-2022.
[10] CoreUI. Free bootstrap admin templates - CoreUI. https://coreui.io/ , 2021. Accessed: 15-Feb-2022.
[11] Himler, Tom; Sharma, Pranav; Taing, Eriq; Akula, Akshay; and Joaquin, Evan. Risk Prediction Sentiment Analysis. 2021. CS4624 submission to VTechWorks, Virginia Tech, Blacksburg, VA 24061 USA. http://hdl.handle.net/10919/107004, accessed May 3, 2022.
[12] Christian Johnson. Risk Insights AI. https://riskinsights.ai/, 2022. Accessed: 15-Feb-2022.
[13] Amazon Web Services. AWS Comprehend. https://aws.amazon.com/comprehend/, 2022. Accessed: 28-Feb-2022.
[14] Amazon Web Services. AWS DynamoDb. https://aws.amazon.com/dynamodb/, 2022. Accessed: 15-Mar-2022.
[15] NodeJs. NodeJs. https://nodejs.org/en/, 2022. Accessed: 20-Feb-2022.
[16] Microsoft. Microsoft Azure Text Analytics. https://azure.microsoft.com/en-us/free/cognitive-services/, 2022. Accessed: 21-Feb-2022.