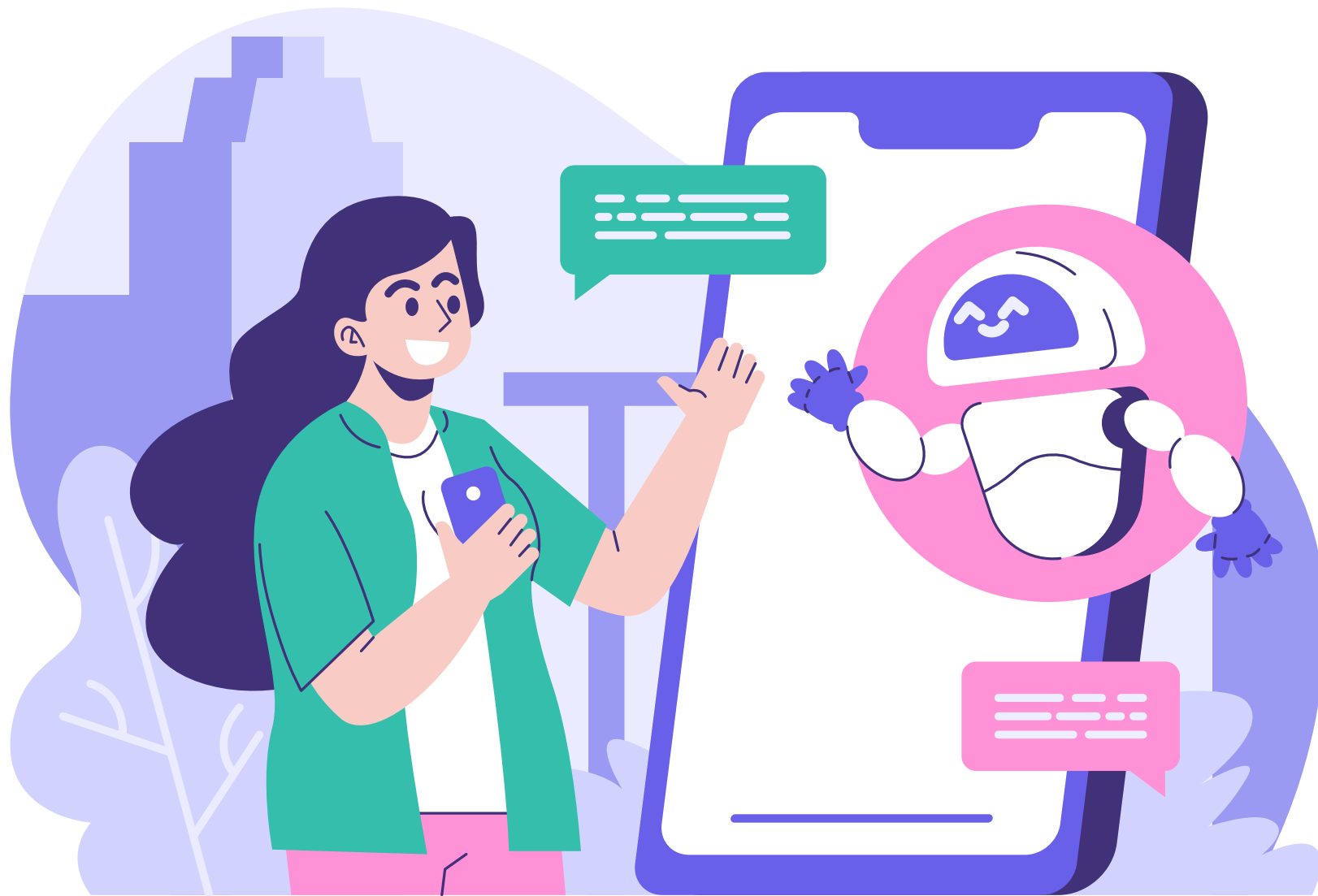


TEAM 7

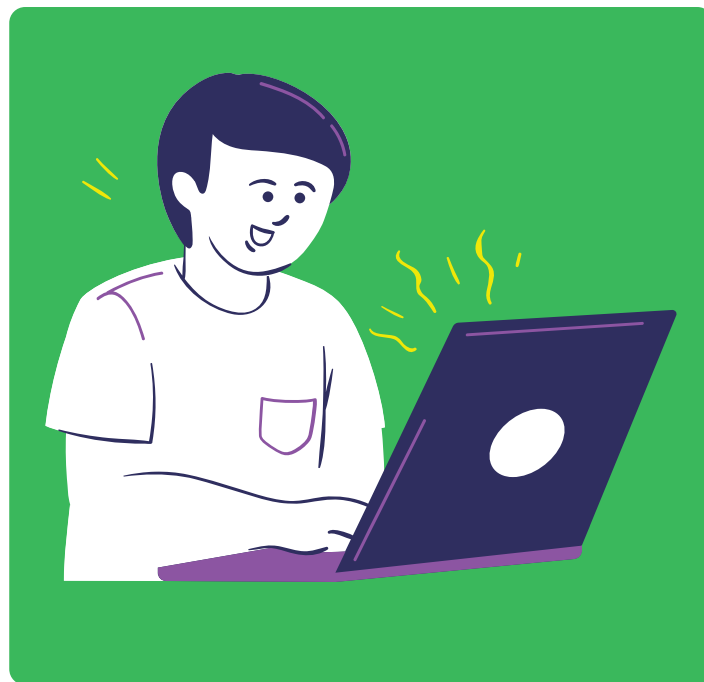
NATURAL LANGUAGE PROCESSING



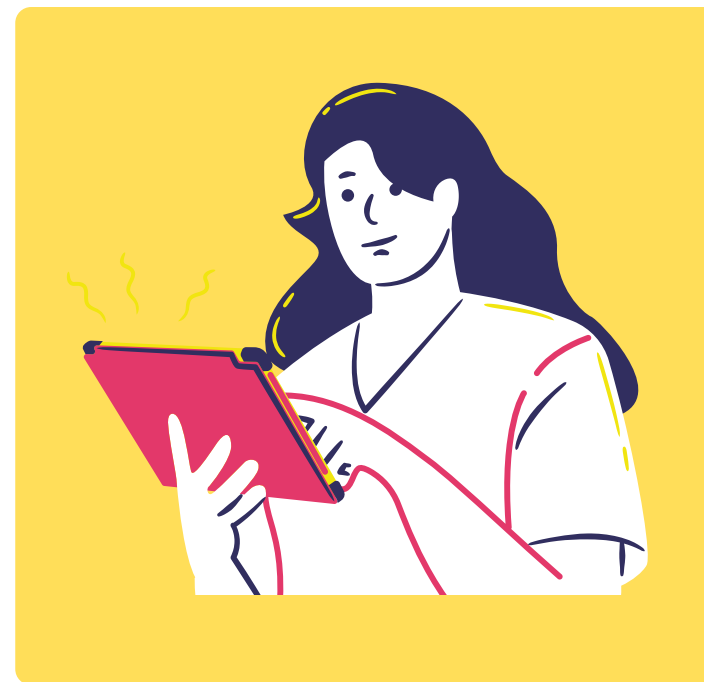
SIOBOT

(Studi Independen Orbit Chatbot)

Our Team



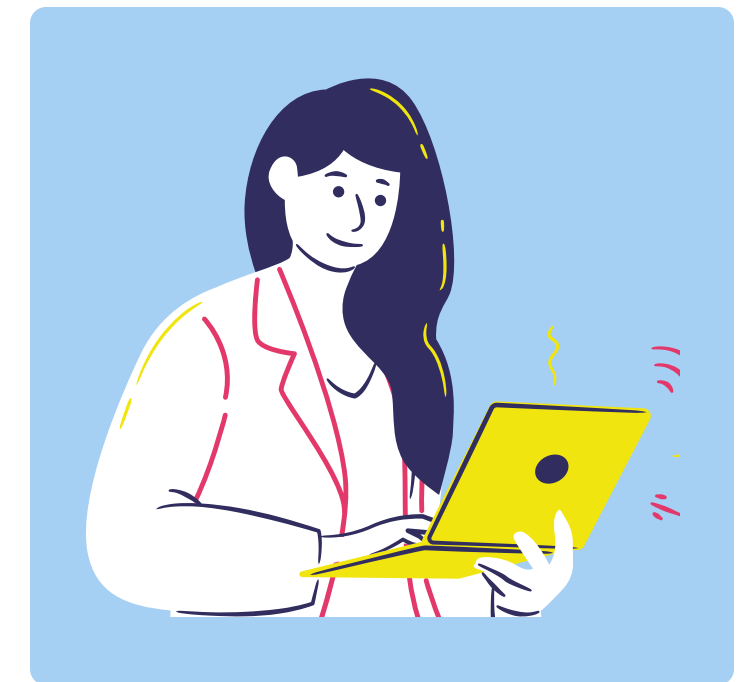
Thomas Marcellino



Nurul Amaliah



Yani Maila Santi



Chilmatus Saidah

Introduction

SIOBOT adalah sebuah aplikasi berbasis web yang nantinya akan membantu calon student Orbit Future Academy untuk membantu menjawab kebingungan nya terkait Studi Independen yang ada di Orbit Future Academy. SIOBOT ini nantinya akan dimunculkan diweb Orbit Future Academy, karena kami melihat masih belum terdapat Chatbot secara live di Web tersebut.



Apa manfaat dari project ini?



01

Membantu mahasiswa yang ingin mendaftar studi independent di Orbit Future Academy

02

Sebagai pusat informasi dari kegiatan Studi Independen Orbit Future Academy

03

Mempercepat proses tanya jawab antara calon mahasiswa dengan pihak Orbit Future Academy

Problem Scoping

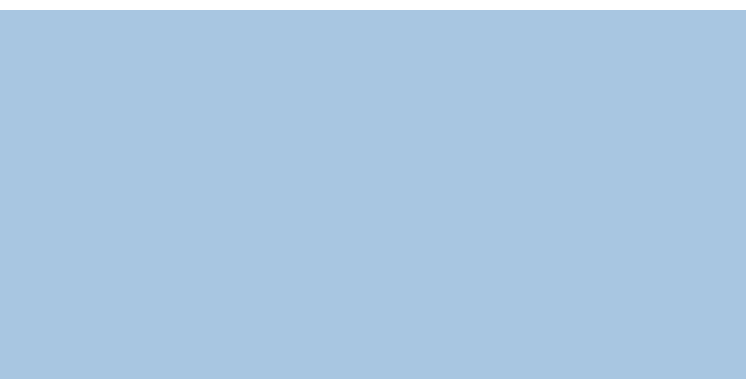
Who ?

Calon student orbit



What ?

Membantu mahasiswa yang akan mengikuti Studi Independen di Orbit Future Academy



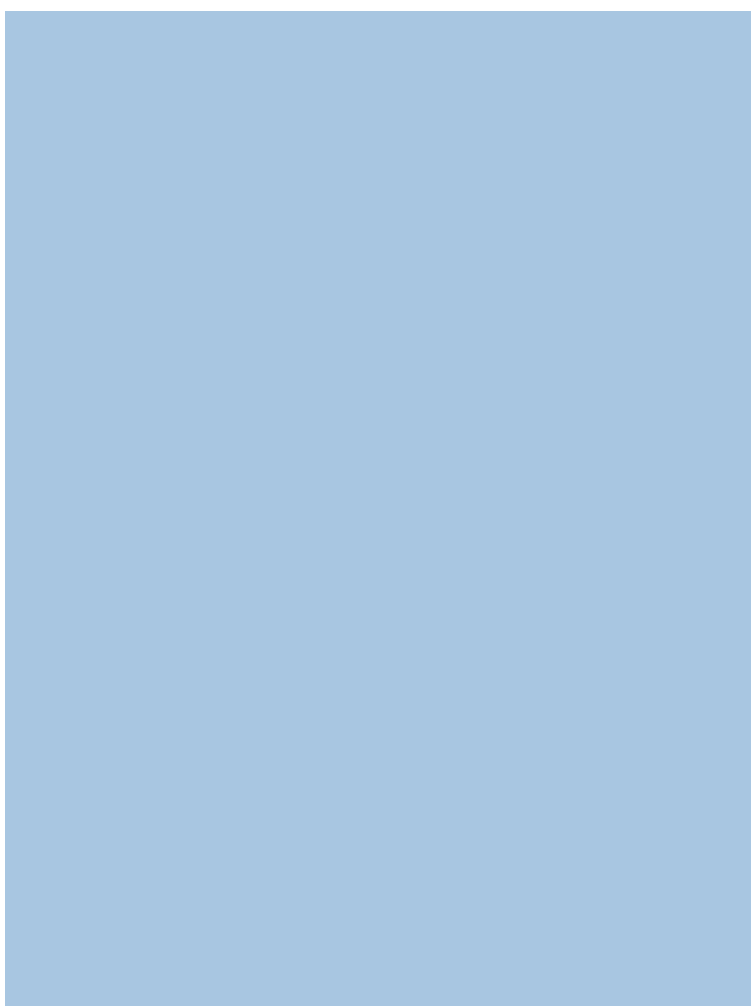
Where ?

Seluruh Wilayah Indonesia



Why ?

Calon student masih mengalami kebingungan akan program studi independen di orbit future academy



Data Acquisition

Pada data acquisition kami menggunakan create data json secara manual

- greeting
- goodbye
- terimakasih
- SIOBOT
- pencipta_siobot
- penjelasan_studi_independen
- persyaratan_studi_independen
- periode_dan_pendaftaran_studi_independen
- manfaat_studi_independen

- penjelasan_orbit
- progam_orbit
- penjelasan_ai4jobs
- penjelasan_aimasteryprogam
- rekomendasi_progam
- pembelajaran_ai4jobs
- pembelajaran_aimasteryprogam
- jam_pembelajaran_orbit



Data Exploration

```

tokenizer = Tokenizer(num_words=2000)
tokenizer.fit_on_texts(data['patterns'])
train = tokenizer.texts_to_sequences(data['patterns'])
train

```

```

[[39],
 [40],
 [41],
 [42],
 [43],
 [44],
 [45],
 [46],
 [47],
 [48],
 [49],
 [50],

```

```

[1, 5, 75, 6, 7],
[31, 76, 77, 78, 79, 80, 81, 6, 7],
[1, 5, 82, 83, 84, 6, 7],
[1, 10, 13, 14, 15],
[12, 13, 14, 15],
[13, 14, 15, 10, 12],
[1, 5, 2, 32, 13, 14, 15, 4, 33, 34, 6, 7],
[2, 1, 5, 4, 33, 34, 6, 7, 11, 13, 14, 15],
[1, 10, 2, 3, 16, 17],
[3, 16, 17, 10, 1],
[12, 3, 16, 17],
[1, 10, 2, 3, 8, 2],
[3, 8, 2, 10, 1],
[12, 3, 8, 2],

```

```

[ ] # Apply padding
x_train = pad_sequences(train)
print(x_train) # Padding Sequences

```

```

[[ 0  0  0  0  0  0  0  0  0  0  0  0 39]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 40]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 41]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 42]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 43]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 44]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 45]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 46]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 47]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 48]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 49]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 50]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 51]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 52]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 53]
 [ 0  0  0  0  0  0  0  0  0  0  1 10 19]
 [ 0  0  0  0  0  0  0  0  0  0  0 20 19]
 [ 0  0  0  0  0  0  0  0  0  0  0 20 54]
 [ 0  0  0  0  0  0  0  0 19  4 55 56 20]
 [ 0  0  0  0  0  0  0  0  0  0 20 57 19]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 58]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 59]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 27]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 60]
 [ 0  0  0  0  0  0  0  0  0  0  0 61 27]
 [ 0  0  0  0  0  0  0  0  0  0  0 62 63]

```

```

lemmatizer = WordNetLemmatizer()
words = [lemmatizer.lemmatize(w.lower()) for w in words if w not in ignore_words]
words = sorted(list(set(words)))

```

```

print (len(words), "unique lemmatized words", words)

```

```

90 unique lemmatized words ['.', '4', 'academy', 'afternoon', 'ai', 'apa', 'apabila', 'apasi', 'bagaimana', 'bai', 'banyak', 'belajar', 'berapa', 'bro', 'bye', 't

```

```

[ ] classes = sorted(list(set(classes)))
print (len(classes), "classes", classes)

```

```

17 classes ['SI0BOT', 'goodbye', 'greeting', 'jam_pembelajaran_orbit', 'manfaat_studi_independen', 'pembelajaran_ai4jobs', 'pembelajaran_aimasteryprogam', 'pencil

```

```

[ ] print (len(documents), "documents")

```

```

422 documents

```


Modelling



Dalam project ini menggunakan bahasa pemrograman Python dengan menggunakan LSTM (Long Short Term Memory). Dimana algoritma Deep Learning ini biasa digunakan untuk membuat prediksi dan klasifikasi yang berhubungan dengan waktu dan data teks. Struktur dari Algoritma ini sendiri terdiri dari Neural Network dan beberapa blok memori yang berbeda. blok memori itni disebut sebagai cell. State dari cell dan hidden state akan diteruskan ke cell berikutnya.

Implementasi Modeling

Dalam project ini menggunakan bahasa pemrograman Python dengan menggunakan LSTM (Long Short Term Memory)

```
# Creating the model (Membuat Modeling)
i = Input(shape=(input_shape,))
x = Embedding(vocabulary+1,10)(i) # Layer Embedding
x = LSTM(10, return_sequences=True)(x) # Layer Long Short Term Memory
x = Flatten()(x) # Layer Flatten
x = Dense(output_length, activation="softmax")(x) # Layer Dense
model = Model(i,x)

# Compiling the model (Kompilasi Model)
model.compile(loss="sparse_categorical_crossentropy", optimizer='adam', metrics=['accuracy'])

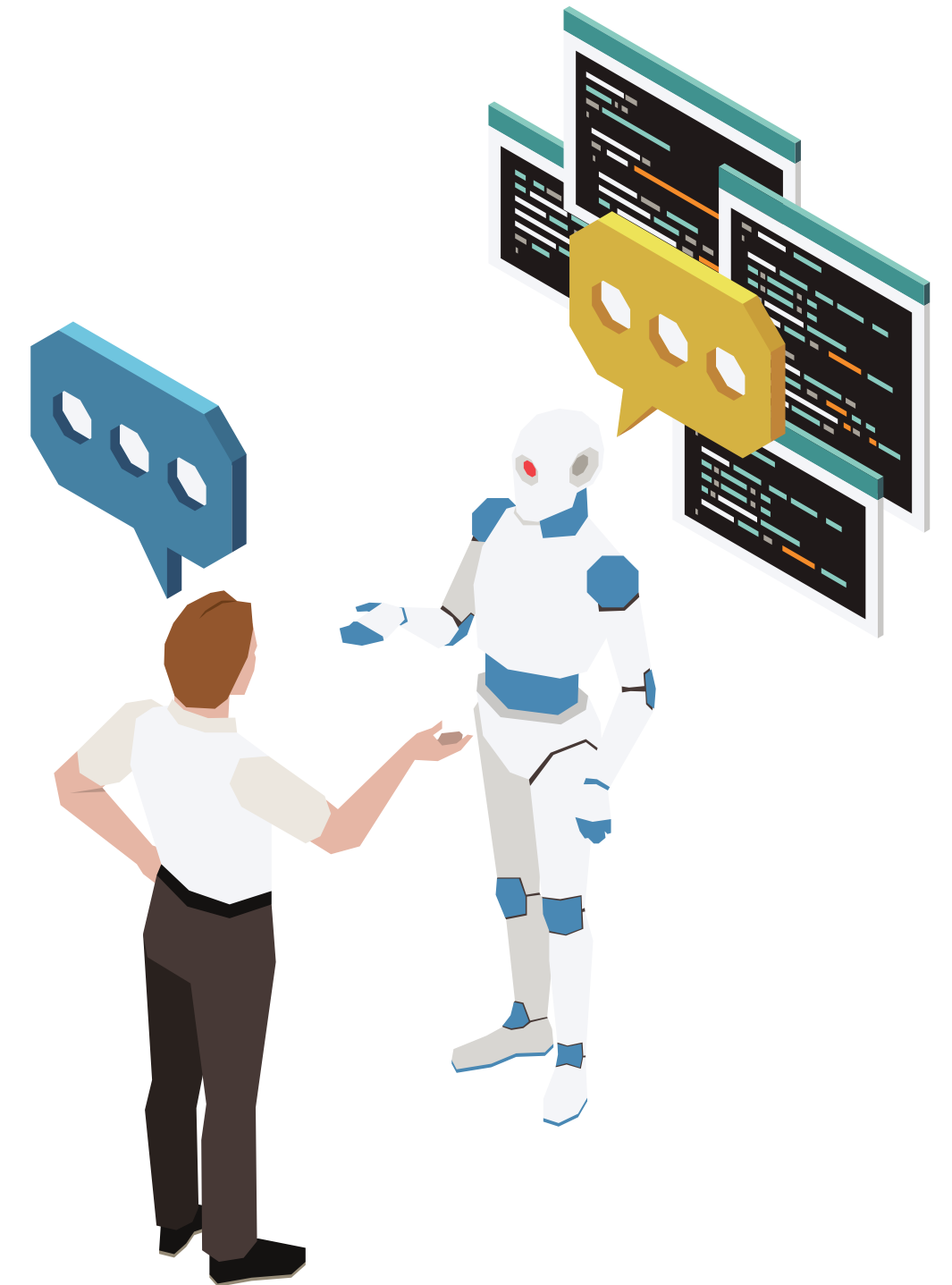
# Visualization Plot Architecture Model (Visualisasi Plot Arsitektur Model)
plot_model(model, to_file='model_plot.png', show_shapes=True, show_layer_names=True)

# Menampilkan Parameter Model
model.summary()

# Training the model (Latih model data sampai 400 kali)
train = model.fit(x_train, y_train, epochs=400)
```

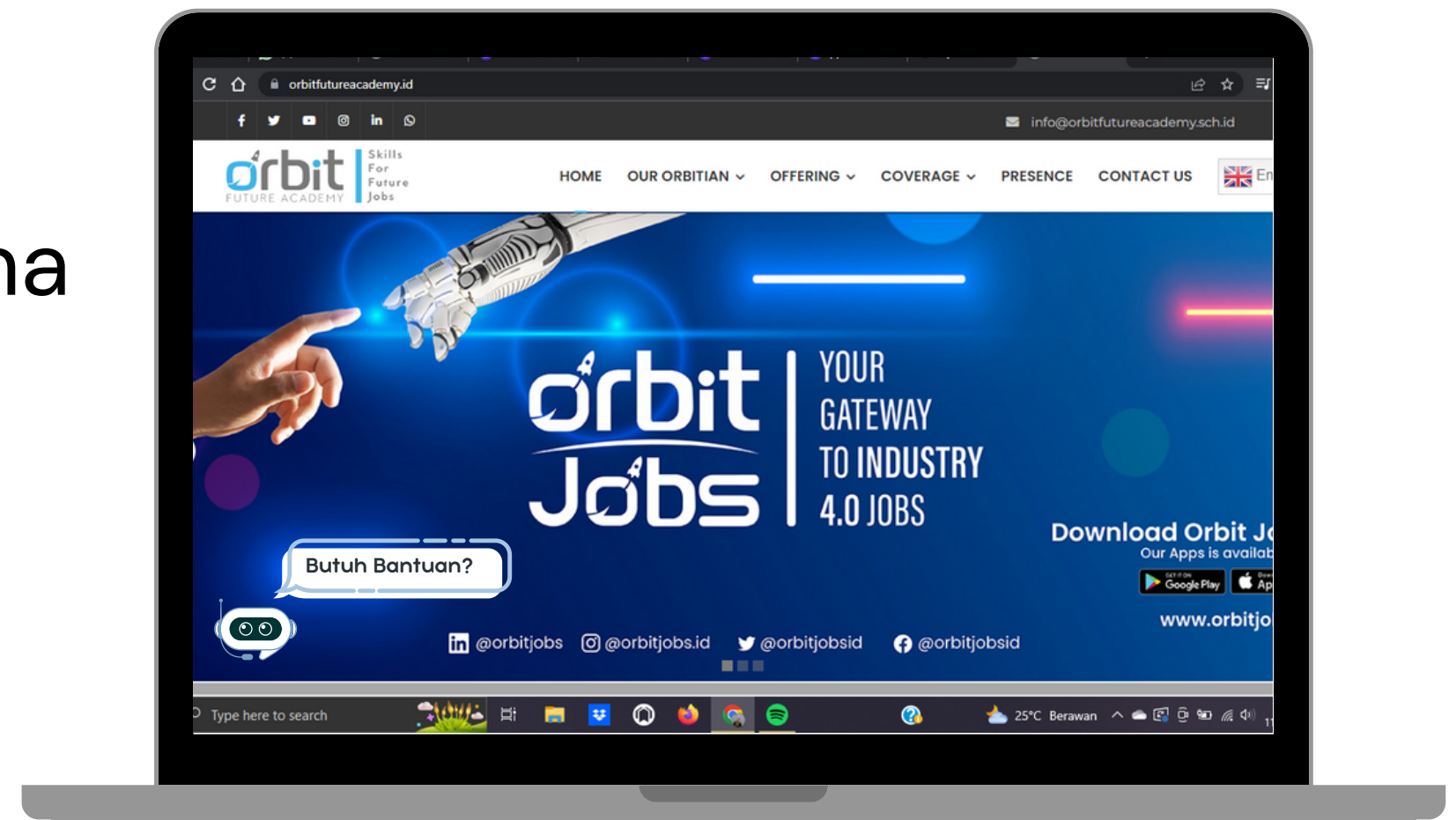
Evaluation

Dengan memasukkan variabel yang digunakan kedalam program, maka akan menghasilkan model chatbot yang dapat membantu user atau penggunaanya dalam mencari informasi serta penjelasan terkait permasalahan yang ditanyakan. Pada project ini, informasi serta penjelasan yang dapat diterima adalah mengenai program Studi Independen pada mitra Orbit Future Academy secara garis besar.



Deployment

SIOBOT ini nantinya akan dijadikan sebagai Aplikasi berbasis web. Dimana hal ini akan mudah sekali digunakan, terutama oleh kalangan calon mahasiswa Studi Independen yang notabennya adalah generasi Z.



TEAM 7

NATURAL LANGUAGE PROCESSING

DEMO



THANK YOU!!!

