

PERSONAL INFORMATION



KEY SKILLS

Sweve	Chauhan
-------	---------

PROFILE SUMMARY

Data Science Enthusiast with bachelor's degree in Economics, I have developed skills in Python, SQL, ML, Data visualization and Statistical Analysis. I am well-versed in working with data preprocessing, model development, and deploying machine learning algorithms to extract actionable insights.

EDUCATION



Projects

python	5 Weeks	C Language identification model
fastapi		• A language identification model is designed to predict the language in which a given text is written.
statistics		• Use dataset <i>language_identification.csv</i> to train the model
natural language processing		 Encode language column into numerical values using LabelEncoder from sklearn
deep learning		Preprocessing function for input texts to remove
machine learning		 Punctuation marks & convert to lowercase CountVectorizer to convert text data into numerical
sql		feature vectors (bag-of-words representation)
docker		Train GaussianNB & MultinoimalNB
		Pipeline is created, combining the CountVectorizer for
Data Visualization		feature extraction and the Multinomial Naive Bayes
Pandas		classifier for classification
Keras		Accuracy, Mean Squared Error (MSE), Confusion Matrix, and Classification Report are computed to evaluate the model's performance
Scikit learn		Trained model is saved to a file using nickle
Numpy		Fastapi is used for apis & deploying on render
	2 Months	Maleria Detection Model
OTHER PERSONAL DETAILS		 Malaria cell images dataset with 27,558 cell images consists of two classes: infected and uninfected Vieweliastion using ou? and methodiib
City Noida		 Visualisation using cv2 and matplotlib ImageDataGenerator is used to load, rescale, preprocess the dataset.

Country INDIA HOBBIES Reading Books, Music, Sports LANGUAGES	 Model Architecture: CNN layer, ReLU activation, MaxPool2D, BatchNormalization & Dropout layer to prevent overfitting. Flatten and Dense layers with Sigmoid activation for binary classification (Infected vs Uninfected) and adam optimizer. Evaluation using validation data, Classification Report and Confusion Matrix Pretrained model : VGG16 is also used Model is saved using Keras and FastAPI is used create API Containerisation using Docker and project is pushed to Github repository Model is deployed on Render Librariae & Modules: teneorflow, Kerae, eklaster, ev2
• english	 Libraries & Modules: tensorflow, Keras, sklearn, cv2, numpy, matplotlib, etc.
• hindi	
	 2 Months Next Word Predictor Goal of code is to train a sequence-to-sequence model that predicts the next word in a sequence. Reading a text file (1661-0.txt) Keras' Tokenizer to convert text into sequences of integer indices One-hot encoded using Keras' to_categorical to transform the integer labels into binary vectors Pad_sequences from Keras to pad the sequences to the same length Two Models are trained : SimpleRNN, LSTMs Model Architecture : Embedding layer, SimpleRNN/LSTMs, SpatialDropout1D and Dropout (prevent overfitting), Batch Normalization (stabilize and accelerate training), Dense layer with softmax function Model is compiled with categorical cross-entropy loss function(multiclass classification problem) and the Adam optimizer model is evaluated using precision, recall, and F1 score Usage of Libraries: Keras/TensorFlow, Pandas, Sickitlearn, Numpy

COURSES & CERTIFICATIONS

• FastAPI

- Docker
- NLP Natural Language Processing with Python