Learn Python from Scratch

The Easiest Way Possible

Vivian Aranha

Welcome to the World of Coding!

Story Time

- Cool idea for a new app or game
- How do you bring it to life?
- Tell a computer exactly what to do, step by step
- That language is **Python**
- Learning how to chat with your computer
- Use words and symbols to get the computer to do awesome stuff for you
- Math, playing music, creating a website, or making a game

Setting Up Your Coding Space

Coding Space

- https://www.python.org
- Why are we doing this?
  - So it understands the magical code you'll write
  - Translator who speaks both "Computer" and "English"
- Picking an Editor
  - Thonny - https://thonny.org
  - Visual Studio Code (VS Code) - https://code.visualstudio.com
  - Google Colab - https://colab.research.google.com

Your First Line of Code

Variables: Your Virtual ‘Boxes’

Talking to the User: *input()*

Math Magic

Decision Time: If, Elif, Else

Loops: Repeating Actions

Lists: Keep Track of Stuff

Functions:
Build Your Own Commands

Mini Project:
Number Guessing Game

Day 1: Welcome Message Generator

Print Statements & "Hello World"

Vivian Aranha

What You'll Learn Today

- What is a print statement?
- How to use print with **text** and **numbers**
- String **formatting** for dynamic messages
- Using **user input** with print
- **Day 1 Project: Welcome Message Generator**

Day 2: Personalized Greeting Program

Variables & Data Types

Vivian Aranha

What You'll Learn Today

- What are Variables?
- **Data Types**: Strings, Integers, Floats, and Booleans
- **Type Conversion**
- **String Formatting**
- **Day 2 Project: Personalized Greeting Program**

Day 3: Simple Calculator

User Input & String Formatting

Vivian Aranha

What You'll Learn Today

- User Input with **input()**
- **Type Conversion** (int() and float())
- String Formatting with **f-Strings**
- Basic **Arithmetic** Operations
- **Day 3 Project: Simple Calculator**

Day 4: Number Comparison Tool

If-Else Statements

Vivian Aranha

What You'll Learn Today

- What are **If-Else Statements**?
- **Comparison** Operators
- **Logical** Operators
- **Nested If-Else** Statements
- **Day 4 Project: Number Comparison Tool**

Day 5: Countdown Timer

Loops (For & While)

Vivian Aranha

What You'll Learn Today

- What are **Loops**?
- **For** Loops
- **While** Loops
- Using **time.sleep()** for Delays
- **Day 5 Project: Countdown Timer**

Day 6: Basic Math Quiz Game

Functions

Vivian Aranha

What You'll Learn Today

- What are **Functions**?
- Defining and **Calling Functions**
- Function **Parameters** and **Arguments**
- **Return** Statements
- **Day 6 Project: Basic Math Quiz Game**

Day 7: Shopping List App

Lists

Vivian Aranha

What You'll Learn Today

- What are **Lists**?
- **List Operations**: Adding, Removing, and Accessing Items
- **Looping** Through Lists
- List **Methods**
- **Day 7 Project: Shopping List App**

Day 8: Contact Book

Dictionaries

Vivian Aranha

What You'll Learn Today

- What are **Dictionaries**?
- **Accessing** and **Modifying** Dictionary Values
- **Adding and Removing Entries**
- **Looping** Through a Dictionary
- **Day 8 Project: Contact Book**

Day 9: Ingredient Checker

Tuples & Sets

Vivian Aranha

What You'll Learn Today

- What are **Tuples**?
- **Tuple Operations and Unpacking**
- What are **Sets**?
- Set Operations (**Union, Intersection, Difference**)
- **Day 9 Project: Ingredient Checker**

Day 10: Note-Taking App

File Handling

Vivian Aranha

What You'll Learn Today

- What is **File Handling**?
- **Reading** from Files
- **Writing** to Files
- **Appending** to Files
- **Day 10 Project: Note-Taking App**

Day 11:
Safe Calculator

Exception Handling

Vivian Aranha

What You'll Learn Today

- What are **Exceptions**?
- Using **try, except, else, and finally**
- Handling **Multiple** Exceptions
- Raising **Custom** Exceptions
- **Day 11 Project: Safe Calculator**

Day 12: Temperature Converter

Functions with Return Values

Vivian Aranha

What You'll Learn Today

- Understanding **Return Values in Functions**
- Using Functions to **Perform Calculations**
- How to **Return Multiple** Values
- **Best Practices** for Return Values
- **Day 12 Project: Temperature Converter**

Day 13: Student Grade Manager

List Comprehensions

Vivian Aranha

What You'll Learn Today

- What are **List Comprehensions**?
- Basic **Syntax and Examples**
- **Filtering** with List Comprehensions
- Using **Conditional** Statements
- **Day 13 Project: Student Grade Manager**

Day 14: Random Password Generator

Modules & Libraries

Vivian Aranha

What You'll Learn Today

- What are **Modules and Libraries**?
- **Importing** Modules
- **Built-in** Python **Libraries**
- Creating and Using **Custom Modules**
- **Day 14 Project: Random Password Generator**

Day 15:
Recipe Viewer App

Reading Files

Vivian Aranha

What You'll Learn Today

- What is **File Reading** in Python?
- **Reading Files** Using **open()**
- **Reading Modes** (r, rb, r+)
- Handling **File Reading Errors**
- **Day 15 Project: Recipe Viewer App**

Day 16:
Daily Journal Logger

Writing Files

Vivian Aranha

What You'll Learn Today

- What is **File Writing** in Python?
- **Writing to Files** (w mode)
- **Appending to Files** (a mode)
- Handling **File Writing Errors**
- **Day 16 Project: Daily Journal Logger**

Day 17: Student Report Generator

CSV Files

Vivian Aranha

What You'll Learn Today

- What are **CSV Files**?
- **Reading CSV** Files
- **Writing to CSV** Files
- Using the **csv Module**
- **Day 17 Project: Student Report Generator**

Day 18:
Mini To-Do App

JSON Files

Vivian Aranha

What You'll Learn Today

- What are **JSON Files**?
- **Reading JSON** Data
- **Writing JSON** Data
- **Modifying JSON** Data
- **Day 18 Project: Mini To-Do App**

Day 19: Weather App using API

APIs (Basics)

Vivian Aranha

What You'll Learn Today

- What is an **API**?
- How **APIs Work**
- Using **API Keys**
- **Fetching** Data from APIs using **requests**
- **Day 19 Project: Weather App using API**

Day 20: Event Countdown Timer

Dates & Time

Vivian Aranha

What You'll Learn Today

- Understanding the **datetime** Module
- Working with **Dates and Times**
- **Formatting** Dates and Times
- Calculating **Time Differences**
- **Day 20 Project: Event Countdown Timer**

Day 21: Wikipedia Article Scraper

Web Scraping

Vivian Aranha

What You'll Learn Today

- What is **Web Scraping**?
- Understanding **HTML Structure**
- Using **requests** to **Fetch Web Pages**
- Using **BeautifulSoup** for **Parsing**
- **Day 21 Project: Wikipedia Article Scraper**

Day 22: Bank Account Simulator

Classes & Objects

Vivian Aranha

What You'll Learn Today

- What are **Classes and Objects**?
- Understanding **Class Attributes** and **Methods**
- **Constructors** (__init__ Method)
- Working with **Multiple Objects**
- **Day 22 Project: Bank Account Simulator**

Day 23: Library Management System

Constructors & Methods

Vivian Aranha

What You'll Learn Today

- What are **Constructors**?
- Using **Instance** Methods
- **Class** Methods vs **Static** Methods
- **Encapsulation** and **Validation**
- **Day 23 Project: Library Management System**

Day 24: Employee Management System

Inheritance

Vivian Aranha

What You'll Learn Today

- What is **Inheritance**?
- **Types** of **Inheritance**
- Using the **super()** Function
- **Method Overriding**
- **Day 24 Project: Employee Management System**

Day 25: Animal Sound Simulator

Polymorphism

Vivian Aranha

What You'll Learn Today

- What is **Polymorphism**?
- **Method Overriding** in **Polymorphism**
- Using **Polymorphism** in Python
- **Real-world Examples** of Polymorphism
- **Day 25 Project: Animal Sound Simulator**

Day 26: Secure User Profile App

Encapsulation

Vivian Aranha

What You'll Learn Today

- What is **Encapsulation**?
- **Public, Protected,** and **Private** Attributes
- **Getter** and **Setter** Methods
- **Validating** User Data
- **Day 26 Project: Secure User Profile App**

Day 27: Inventory Management System

Static & Class Methods

Vivian Aranha

What You'll Learn Today

- What are **Static and Class Methods**?
- When to Use **Static vs Class Methods**
- Defining and **Calling Static** & **Class Methods**
- Practical **Use Cases**
- **Day 27 Project: Inventory Management System**

Day 28:
Mini ATM Machine

Final OOP Project

Vivian Aranha

What You'll Learn Today

- Combining **OOP Principles** in a **Real-World Project**
- Designing a **Modular Class Structure**
- Implementing **Secure User Authentication**
- Managing **Account Transactions**
- **Day 28 Project: Mini ATM Machine**

Day 29:
Simple GUI App

Tkinter Basics

Vivian Aranha

What You'll Learn Today

- What is **Tkinter**?
- Creating a **Basic Tkinter Window**
- Adding **Widgets** (Labels, Buttons, Entry Fields)
- Handling **User Events**
- **Day 29 Project: Simple GUI App**

Day 30:
Click Counter App

Buttons & Events

Vivian Aranha

What You'll Learn Today

- **How Buttons Work** in Tkinter
- **Binding Events** to Buttons
- **Dynamic Updates** Using Button Events
- Managing **Button States**
- **Day 30 Project: Click Counter App**

Day 31:
BMI Calculator

Input Fields

Vivian Aranha

What You'll Learn Today

- Understanding **Input Fields in Tkinter**
- Getting and **Validating User Input**
- Displaying **Dynamic Results**
- Using **Entry Widgets** with **Labels and Buttons**
- **Day 31 Project: BMI Calculator**

Day 32:
Drawing Pad App

Canvas Widgets

Vivian Aranha

What You'll Learn Today

- What is a **Canvas Widget**?
- Drawing **Shapes and Lines**
- Handling **Mouse Events** on Canvas
- **Clearing and Resetting** the Canvas
- **Day 32 Project: Drawing Pad App**

Day 33:
Simple Login System

Message Boxes

Vivian Aranha

What You'll Learn Today

- What are **Message Boxes**?
- **Types of Message Boxes**
- Using **Message Boxes** for Validation
- Handling **User Authentication**
- **Day 33 Project: Simple Login System**

Day 34:
To-Do List GUI

Advanced Widgets

Vivian Aranha

What You'll Learn Today

- Introduction to **Advanced Tkinter Widgets**
- Using **Listbox** for **Dynamic Lists**
- **Scrollbar Integration**
- Handling **User Actions** (Add, Delete, Clear)
- **Day 34 Project: To-Do List GUI**

Day 35:
Expense Tracker App

GUI Capstone

Vivian Aranha

What You'll Learn Today

- Combining **Tkinter Widgets** for a Complex App
- Managing **User Input and Validation**
- Displaying **Dynamic Data in a Listbox**
- Implementing **File Handling for Data Persistence**
- **Day 35 Capstone Project: Expense Tracker App**

Day 36:
Hello Flask App

Flask Basics

Vivian Aranha

What You'll Learn Today

- What is **Flask**?
- **Setting Up Flask**
- Creating Your **First Flask Route**
- Understanding **Flask Templates**
- **Day 36 Project: Hello Flask App**

Day 37: Personal Blog Website

Routes & Templates

Vivian Aranha

What You'll Learn Today

- **Advanced** Flask Routing
- **Dynamic** Templates with **Jinja2**
- **Passing Data** Between **Routes** and **Templates**
- Organizing **Flask Projects for Scalability**
- **Day 37 Project: Personal Blog Website**

Day 38: Contact Form App

Forms & User Input

Vivian Aranha

What You'll Learn Today

- Introduction to **Flask Forms**
- **Installing Flask-WTF**
- Creating Forms with **Flask-WTF**
- **Validating** User Input
- **Day 38 Project: Contact Form App**

Day 39: User Registration App

Database Integration

Vivian Aranha

What You'll Learn Today

- Introduction to **Flask-SQLAlchemy**
- **Setting Up** the Database
- Creating **Models for User Data**
- Handling **User Registration**
- **Day 39 Project: User Registration App**

Day 40: Mini Weather API

REST APIs

Vivian Aranha

What You'll Learn Today

- What are **REST APIs**?
- Setting Up a **Flask REST API**
- Defining **API Routes and Endpoints**
- Returning **JSON Responses**
- **Day 40 Project: Mini Weather API**

Day 41: Deploy Flask App

Deployment

Vivian Aranha

What You'll Learn Today

- Introduction to **Flask Deployment**
- Preparing Your **Flask App for Deployment**
- Deployment to **Heroku**
- **Environment Variables** and **Secrets Management**
- **Day 41 Project: Deploy Flask App on Heroku**

Day 42: Portfolio Website

Flask Capstone

Vivian Aranha

What You'll Learn Today

- Structuring a **Flask Portfolio Website**
- Creating **Dynamic Routes and Templates**
- Adding a **Contact Form**
- **Database** for Projects
- **Day 42 Project: Portfolio Website**

Day 43:
Matrix Calculator

NumPy

Vivian Aranha

What You'll Learn Today

- Introduction to **NumPy**
- **Matrix Operations** with NumPy
- Handling **User Input for Matrices**
- Building a **Matrix Calculator**
- **Day 43 Project: Matrix Calculator**

Day 44:
Data Cleaner

Pandas

Vivian Aranha

What You'll Learn Today

- Introduction to **Pandas**
- **Loading** and **Inspecting** Data
- **Data Cleaning Techniques**
- **Transforming** and **Exporting** Clean Data
- **Day 44 Project: Data Cleaner**

Day 45:
Graph Plotter

Matplotlib

Vivian Aranha

What You'll Learn Today

- Introduction to **Matplotlib**
- Creating **Basic Plots**
- **Customizing** Graphs
- Plotting from **Data Files**
- **Day 45 Project: Graph Plotter**

Day 46:
Sales Report Analyzer

Data Analysis

Vivian Aranha

What You'll Learn Today

- Introduction to **Data Analysis**
- **Loading** and **Exploring Sales** Data
- **Cleaning** and **Transforming** Data
- Generating **Insights** and **Visualizations**
- **Day 46 Project: Sales Report Analyzer**

Day 47: Temperature Plotter

Plotting Trends

Vivian Aranha

What You'll Learn Today

- **Loading and Processing** Temperature Data
- Plotting **Temperature Trends**
- Highlighting **Anomalies** and **Averages**
- **Customizing** and **Saving** Plots
- **Day 47 Project: Temperature Plotter**

Day 48:
Stock Price Tracker

Data Scraping

Vivian Aranha

What You'll Learn Today

- Understanding **Web Scraping Basics**
- Using **requests** and **BeautifulSoup** for Scraping
- **Extracting Stock Prices** from a Website
- Adding **Dynamic Updates for Real-Time Tracking**
- **Day 48 Project: Stock Price Tracker**

Day 49: Global Weather Dashboard

Capstone

Vivian Aranha

What You'll Learn Today

- Fetching **Real-Time Weather Data**
- Visualizing **Weather Trends**
- Creating a **User-Friendly Dashboard**
- Enhancing Functionality with **Advanced Features**
- **Day 49 Capstone Project: Global Weather Dashboard**

Day 50: Weather Dashboard App

Vivian Aranha

What You'll Learn Today

- Using a **Weather API to Fetch Data**
- **Processing and Displaying** Weather Information
- Creating a **Clean Web Interface with Flask**
- Enhancing the Dashboard with **Interactive Features**
- **Day 50 Project: Weather Dashboard App**

Day 51:
Expense Tracker

Vivian Aranha

What You'll Learn Today

- **Logging** and **Storing** Expenses
- **Categorizing** and **Summarizing** Data
- Visualizing **Expenses** with **Graphs**
- Building a **User-Friendly Interface**
- **Day 51 Project: Expense Tracker App**

Day 52:
File Organizer Tool

Vivian Aranha

What You'll Learn Today

- **Identifying File Types** and **Categories**
- **Moving Files** into Folders
- Automating **File Organization**
- Adding **Customization Options**
- **Day 52 Project: File Organizer Tool**

Day 53:
Tic-Tac-Toe Game

Vivian Aranha

What You'll Learn Today

- Designing the **Game Logic**
- Building the **User Interface with Tkinter**
- Handling **Player Turns** and **Game Status**
- Enhancing the **User Experience**
- **Day 53 Project: Tic-Tac-Toe Game**

Day 54:
Mini Chatbot

Vivian Aranha

What You'll Learn Today

- Understanding **Chatbot Basics**
- Handling **User Input and Responses**
- Adding **Basic Conversation Logic**
- Improving the **Chatbot with Custom Responses**
- **Day 54 Project: Mini Chatbot**

Day 55: Music Playlist Organizer

Vivian Aranha

What You'll Learn Today

- **File and Directory Management** using Python's **os** and **shutil** modules
- Metadata Extraction using libraries like **mutagen**
- **Automating Organization** based on **extracted metadata**
- Exporting **Summaries to JSON or CSV** for analysis
- **Day 55 Project: Music Playlist Organizer**

Day 56: Personal Budget Planner

Vivian Aranha

What You'll Learn Today

- Setting Up **Budget Categories** and **Goals**
- Tracking **Income and Expenses**
- Calculating **Savings and Insights**
- Visualizing **Financial Data**
- **Day 56 Project: Personal Budget Planner**

Day 57:
ASCII Art Generator

Vivian Aranha

What You'll Learn Today

- Understanding **ASCII Art Basics**
- **Loading** and **Processing Images**
- Mapping **Pixels to ASCII** Characters
- Generating **ASCII Art**
- **Day 57 Project: ASCII Art Generator**

Day 58:
Pomodoro Timer

Vivian Aranha

What You'll Learn Today

- Understanding the **Pomodoro Technique**
- Creating the **Timer Logic**
- Building the **GUI with Tkinter**
- Adding **Alerts** and **Customizations**
- **Day 58 Project: Pomodoro Timer**

Day 59: Markdown to HTML Converter

Vivian Aranha

What You'll Learn Today

- Understanding **Markdown and HTML Basics**
- Parsing **Markdown Syntax**
- Generating **HTML Content**
- Adding **Custom Styles**
- **Day 59 Project: Markdown to HTML Converter**

Day 60:
Personal Diary App

Vivian Aranha

What You'll Learn Today

- Setting Up a **Secure Diary System**
- **Encrypting** Diary Entries
- **Creating, Viewing,** and **Managing** Entries
- Adding **Password Protection**
- **Day 60 Project: Personal Diary App**

Day 61:
Social Media Scraper

Vivian Aranha

What You'll Learn Today

- Understanding **HTML Structure and Tags**
- Using Python to **Parse HTML Files**
- **Extracting** Target Information
- **Formatting** and **Saving** Extracted Data
- **Day 61 Project: Social Media Scraper**

Day 62: Automated Backup Tool

Vivian Aranha

What You'll Learn Today

- Handling **File Operations in Python**
- **Creating** and **Managing** Backup Directories
- **Automating** the Backup Process with **Timestamps**
- Generating **Logs for Backup Activities**
- **Day 62 Project: Automated Backup Tool**

Day 63: Movie Recommendation System

Vivian Aranha

What You'll Learn Today

- Understanding **Collaborative Filtering**
- Working with a **Movie Dataset**
- Building the **Recommendation System**
- Computing **Cosine Similarity**
- **Day 63 Project: Movie Recommendation System**

Day 64:
PDF Merger Tool

Vivian Aranha

What You'll Learn Today

- Understanding **PDF File Handling**
- Using Python for **Merging PDFs**
- Adding **File Order** Options
- Implementing a **User-Friendly CLI**
- **Day 64 Project: PDF Merger Tool**

Day 65: Portfolio Website Backend

Vivian Aranha

What You'll Learn Today

- Serving **Static Files**
- Creating a **Flask App**
- Handling **User Feedback with Forms**
- Storing **Feedback Locally**
- **Day 65 Project: Portfolio Website Backend**

Day 66: Flashcards Learning App

Vivian Aranha

What You'll Learn Today

- Creating a **Flashcard Data Structure**
- **Adding** and **Reviewing** Flashcards
- **Marking** Cards as **Learned**
- Implementing the **CLI**
- **Day 66 Project: Flashcards Learning App**

Day 67: Stock Market Dashboard

Vivian Aranha

What You'll Learn Today

- Working with **Stock Price Data**
- Creating a **Tkinter Dashboard**
- **Plotting Stock Trends** with Matplotlib
- Adding **Interactivity** to the Dashboard
- **Day 67 Project: Stock Market Dashboard**

Day 68:
Task Scheduler

Vivian Aranha

What You'll Learn Today

- Setting Up **Task Data Management**
- Creating an **Interactive GUI with Tkinter**
- Adding **Reminder** and **Completion** Features
- Adding **Save** and **Delete** Features
- **Day 68 Project: Task Scheduler**

Day 69:
Currency Converter

Vivian Aranha

What You'll Learn Today

- Setting Up **Exchange Rate Data**
- Building the **Conversion Logic**
- Creating a **GUI for Interactivity**
- Validating **User Input**
- **Day 69 Project: Currency Converter**

Day 70:
Data Visualizer App

Vivian Aranha

What You'll Learn Today

- **Reading** and **Handling** CSV/Excel Files
- Creating Interactive **GUI with Tkinter**
- **Dynamic Data Visualization** Using Matplotlib
- **Error Handling** for Data Input
- **Day 70 Project: Data Visualizer App**

Day 71:
Spam Email Detector

Vivian Aranha

What You'll Learn Today

- Understanding **Spam Detection** and **NLP**
- **Preprocessing** Text Data
- Training a **Machine Learning Model**
- Evaluating **Model Performance**
- **Day 71 Project: Spam Email Detector**

Day 72: Text Sentiment Analyzer

Vivian Aranha

What You'll Learn Today

- Understanding **Sentiment Analysis**
- Using **Pre-Trained NLP Models** for Sentiment Detection
- Building a **Sentiment Analyzer** Using **TextBlob** and **VADER**
- Adding **User Input** for **Real-Time** Analysis
- **Day 72 Project: Text Sentiment Analyzer**

Day 73: Handwriting Digit Recognition

Vivian Aranha

What You'll Learn Today

- Understanding **Image Classification** & **Neural Networks**
- Loading and Preprocessing the **MNIST Dataset**
- Building a **Convolutional Neural Network (CNN)** Model
- **Training** & **Evaluating** the Model
- Testing with **Custom Handwritten Digits**
- **Day 73 Project: Handwriting Digit Recognition**

Day 74:
Voice Assistant

Vivian Aranha

What You'll Learn Today

- Using **Speech Recognition** to Process **Voice Commands**
- Converting **Text to Speech (TTS)** Output
- Implementing **Basic Commands** (e.g., time, weather, Wikipedia search)
- Building an **Interactive Virtual Assistant**
- **Day 74 Project: Voice Assistant**

Day 75:
Face Detection App

Vivian Aranha

What You'll Learn Today

- Understanding **Face Detection** and **OpenCV**
- Setting Up **OpenCV** and Loading **Pre-Trained Face Detection Models**
- Detecting **Faces in Images**
- **Real-Time** Face Detection from **Webcam**
- **Day 75 Project: Face Detection App**

Day 76:
Simple Recommendation System

Vivian Aranha

What You'll Learn Today

- Understanding **Recommendation Systems**
- Building a **Content-Based Recommendation System**
- Implementing **Collaborative Filtering**
- **Day 76 Project: Simple Recommendation System**

Day 77:
AI Chatbot with NLP

Vivian Aranha

What You'll Learn Today

- Understanding **NLP-Based Chatbots**
- Preprocessing Text Data for **Chatbot Training**
- Using **Rule-Based** and **Machine Learning** Approaches
- Building an **AI-Powered Chatbot** with **Transformers**
- **Day 77 Project: AI Chatbot with NLP**

Day 78:
Object Detection App

Vivian Aranha

What You'll Learn Today

- Understanding **Object Detection** and **Deep Learning**
- Using Pre-Trained **YOLOv5** and **MobileNet SSD** Models
- Processing Images for **Object Detection**
- **Real-Time** Object Detection from **Webcam**
- **Day 78 Project: Object Detection App**

Day 79: Language Translator Tool

Vivian Aranha

What You'll Learn Today

- Understanding **Language Translation APIs**
- Using **Google Translate API** with **Python**
- Building a **Translator App** with User Input
- Creating a **GUI-Based Translator** (Bonus)
- **Day 79 Project: Language Translator Tool**

Day 80:
Fake News Detector

Vivian Aranha

What You'll Learn Today

- Understanding **Fake News Detection** with **Machine Learning**
- Data **Preprocessing** and **Feature Extraction**
- Training a **Machine Learning Model** to Detect Fake News
- Evaluating **Model Performance**
- **Day 80 Project: Fake News Detector**

Machine Learning Algorithms & Implementation

...

Vivian Aranha

Types of Algorithms

- Supervised Learning Algorithms
- Unsupervised Learning Algorithms
- Other specialized categories

Supervised Learning Algorithms

- Regression Algorithms
  - Linear Regression
  - Ridge and Lasso Regression
  - Polynomial Regression
- Classification Algorithms
  - Logistic Regression
  - K-Nearest Neighbors (KNN)
  - Support Vector Machines (SVM)
  - Decision Trees
  - Random Forests
  - Gradient Boosting
  - Naive Bayes

Unsupervised Learning Algorithms

- Clustering Algorithms
  - K-Means Clustering
  - Hierarchical Clustering
  - DBSCAN (Density-Based Spatial Clustering of Applications with Noise)
  - Gaussian Mixture Models (GMM)
- Dimensionality Reduction Algorithms
  - Principal Component Analysis (PCA)
  - t-Distributed Stochastic Neighbor Embedding (t-SNE)
  - Autoencoders

Other Specialized Categories

- Semi-Supervised Learning
  - Self-Training
- Reinforcement Learning
  - Q-Learning
  - Deep Q-Networks (DQN)
  - Policy Gradient Methods
- Anomaly Detection Algorithms
  - One-Class SVM
  - Isolation Forest
- Neural Networks (Deep Learning)
  - Convolutional Neural Networks (CNNs)
  - Recurrent Neural Networks (RNNs)
  - Long Short-Term Memory (LSTM)
  - Transformers

Linear Regression

...

Vivian Aranha

Linear Regression

Linear Regression is a supervised learning algorithm used for predicting a continuous target variable based on one or more input features. It finds the line of best fit (linear relationship) by minimizing the sum of squared di8erences between the actual and predicted values.

Ridge and Lasso Regression

...

Vivian Aranha

Ridge and Lasso Regression

Ridge and Lasso Regression are regularization techniques applied to Linear Regression to prevent overfitting by penalizing large coefficients:

- **Ridge Regression** adds an L2 penalty (sum of squared coefficients).
- **Lasso Regression** adds an L1 penalty (sum of absolute values of coefficients), which can lead to feature selection by shrinking some coefficients to zero.

Polynomial Regression

...

Vivian Aranha

Polynomial Regression

Polynomial Regression is an extension of Linear Regression that models the relationship between the input features and the target variable as an nth-degree polynomial. It can capture non-linear relationships in the data by adding polynomial terms to the features.

Logistic Regression

...

Vivian Aranha

Logistic Regression

Logistic Regression is a supervised learning algorithm used for binary classification problems (e.g., yes/no, spam/not spam). Instead of predicting a continuous output, it predicts the probability of an observation belonging to a particular class by applying the logistic (sigmoid) function, which outputs values between 0 and 1.

K-Nearest Neighbors (KNN)

...

Vivian Aranha

K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) is a simple, non-parametric classification (or regression) algorithm. It classifies new data points based on the majority class of the k-nearest points in the feature space. It's particularly useful for smaller datasets where the relationships among data points can be easily visualized.

Support Vector Machines (SVM)

...

Vivian Aranha

Support Vector Machines (SVM)

Support Vector Machines (SVM) is a powerful classification algorithm that works by finding the hyperplane that best separates classes in the feature space. SVM aims to maximize the margin between the classes, making it a good choice for binary classification, especially when classes are well-separated.

Decision Trees

...

Vivian Aranha

Decision Trees

Decision Trees are a versatile supervised learning algorithm used for both classification and regression. They work by recursively splitting the data into subsets based on the feature that provides the most information gain. Each node represents a decision based on a feature, and each leaf node represents a prediction.

Random Forests

Vivian Aranha

Random Forests

Random Forests are an ensemble learning method that combines multiple decision trees to make a more accurate and stable prediction. Each tree in the forest is trained on a random subset of the data, and the final prediction is made by averaging (for regression) or voting (for classification) the predictions of individual trees. This helps to reduce overfitting and improve generalization.

Gradient Boosting

...

Vivian Aranha

Gradient Boosting

Gradient Boosting is an ensemble technique that builds a series of decision trees, where each tree corrects the errors of the previous ones. By combining the predictions of these trees, Gradient Boosting models create a more accurate final prediction. Popular implementations include XGBoost, LightGBM, and CatBoost, which are optimized for speed and accuracy.

Naive Bayes

...

Vivian Aranha

Naive Bayes

Naive Bayes is a probabilistic classifier based on Bayes' theorem, which assumes that the features are conditionally independent given the class label. Despite this “naive” assumption, it often performs well in text classification and spam detection tasks.

K-Means Clustering

...

Vivian Aranha

K-Means Clustering

K-Means Clustering is an unsupervised learning algorithm that partitions data into k clusters. Each cluster is defined by its centroid, and each data point is assigned to the nearest cluster. The algorithm iteratively adjusts centroids to minimize the variance within each cluster.

Hierarchical Clustering

...

Vivian Aranha

Hierarchical Clustering

Hierarchical Clustering is an unsupervised learning algorithm that builds a hierarchy of clusters. It starts with each data point as its own cluster and then merges or splits clusters based on distance measures, forming a tree-like structure called a dendrogram. The hierarchy can be used to choose a suitable number of clusters by "cutting" the tree at a specific level.

DBSCAN (Density-Based Spatial Clustering of Applications with Noise)

...

Vivian Aranha

DBSCAN

DBSCAN is an unsupervised clustering algorithm that groups data points based on density, making it particularly effective for identifying clusters of arbitrary shapes and for handling noise (outliers). DBSCAN requires two parameters: **eps** (the maximum distance between two points to be considered neighbors) and **min_samples** (the minimum number of points required to form a dense region).

Gaussian Mixture Models (GMM)

...

Vivian Aranha

Gaussian Mixture Models (GMM)

Gaussian Mixture Models (GMM) is a probabilistic clustering algorithm that assumes data points are generated from a mixture of several Gaussian distributions with unknown parameters. GMM assigns a probability to each data point for belonging to each cluster, making it a soft clustering technique. It is particularly useful when clusters have different shapes or densities.

Principal Component Analysis (PCA)

...

Vivian Aranha

Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a dimensionality reduction technique used to transform a high-dimensional dataset into a lower-dimensional one by identifying the directions (principal components) that capture the maximum variance in the data. PCA is widely used for data visualization, noise reduction, and speeding up machine learning algorithms by reducing the number of features.

t-Distributed Stochastic Neighbor Embedding (t-SNE)

...

Vivian Aranha

t-Distributed Stochastic Neighbor Embedding (t-SNE)

t-SNE is a dimensionality reduction technique primarily used for visualizing high-dimensional data in 2D or 3D space. Unlike PCA, t-SNE is non-linear and focuses on preserving the local structure of data, making it highly effective for visualizing clusters. However, it is computationally intensive and best suited for small to medium-sized datasets.

Autoencoders

Vivian Aranha

Autoencoders

Autoencoders are neural networks used for unsupervised learning, specifically for dimensionality reduction and feature extraction. They work by encoding input data into a compressed (latent) representation and then reconstructing the original input from this representation. Autoencoders are useful for tasks like denoising, anomaly detection, and pretraining for other neural networks.

Self-Training

...

Vivian Aranha

Self-Training

Self-Training is a semi-supervised learning approach that leverages a small labeled dataset alongside a larger unlabeled dataset. The model is initially trained on labeled data, and then it makes predictions on the unlabeled data. The confident predictions (those with high certainty) are then added to the labeled dataset, and the process is repeated to improve the model.

Q-Learning

...

Vivian Aranha

Q-Learning

Q-Learning is a model-free reinforcement learning algorithm used to find the optimal action-selection policy for a given problem. It learns by interacting with an environment, updating a Q-table (a matrix of state-action values), and maximizing the expected cumulative reward. Q-Learning is effective in problems where the environment can be represented by discrete states and actions.

Deep Q-Networks (DQN)

•••

Vivian Aranha

Deep Q-Networks (DQN)

Deep Q-Networks (DQN) is a reinforcement learning algorithm that combines Q-Learning with deep neural networks. It uses a neural network to approximate the Q-values for each action in a given state, allowing it to handle environments with high-dimensional and continuous state spaces. DQN uses experience replay (storing past experiences and training on random batches) and a target network to stabilize training.

Policy Gradient Methods

...

Vivian Aranha

Policy Gradient Methods

Policy Gradient Methods are a class of reinforcement learning algorithms that learn a policy directly by optimizing the parameters of a policy network. Instead of learning Q-values like Q-Learning or DQN, policy gradient methods focus on finding the optimal action-selection strategy that maximizes cumulative rewards. A popular approach is the **REINFORCE algorithm**, where actions are sampled from a policy distribution, and the policy is updated using gradients based on rewards.

One-Class SVM

...

Vivian Aranha

One-Class SVM

One-Class SVM (Support Vector Machine) is an algorithm for anomaly detection that identifies data points that differ significantly from the normal distribution of data. It's particularly useful when the dataset primarily consists of one class, and we want to detect outliers. One-Class SVM separates the data into a high-density region (normal data) and sparse regions (anomalies).

Isolation Forest

...

Vivian Aranha

Isolation Forest

Isolation Forest is an ensemble method for anomaly detection that isolates anomalies rather than profiling normal data. The algorithm randomly selects a feature and a split value to partition the data, creating trees where anomalies are easier to isolate due to their sparse distribution. Anomalies are identified based on their shorter path lengths in the tree structure, as they are isolated faster than normal points.

Convolutional Neural Networks (CNNs)

...

Vivian Aranha

Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are deep learning models specifically designed for processing structured grid data, such as images. CNNs use convolutional layers that apply filters to the input image, capturing spatial hierarchies and features like edges, textures, and shapes. CNNs are widely used in computer vision tasks like image classification, object detection, and segmentation.

Recurrent Neural Networks (RNNs)

...

Vivian Aranha

Recurrent Neural Networks (RNNs)

Recurrent Neural Networks (RNNs) are neural networks designed for sequential data, such as time series, language, or speech. RNNs have connections that form cycles, allowing them to retain information from previous steps in the sequence. This makes RNNs well-suited for tasks like text generation, language modeling, and time series forecasting. A common variant, Long Short-Term Memory (LSTM), helps to address the issue of long-term dependency.

Long Short-Term Memory (LSTM)

...

Vivian Aranha

Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) networks are a type of RNN specifically designed to capture long-term dependencies in sequential data. LSTMs use gating mechanisms to control the flow of information, which helps prevent the vanishing gradient problem that standard RNNs suffer from. They are commonly used in tasks like language modeling, machine translation, and time series prediction.

Transformers

...

Vivian Aranha

Transformers

Transformers are deep learning architectures designed for handling sequential data without relying on recurrence, which is commonly used in RNNs. Instead, Transformers use a mechanism called **self-attention** to process all tokens in the sequence simultaneously, capturing dependencies between tokens regardless of their distance in the sequence.
Transformers have become the foundation of many NLP tasks and models, including BERT and GPT.