MACHINE LEARNING INTERNSHIP TASKS

30 Days

CODEVO SOLUTIONS

hr@codevo.in | www.codevo.in

TASKS

Task 1: Wine Quality Prediction (Beginner Level).

Task 2: Sales Forecasting (Beginner Level).

Task 3: Detecting Spam Emails Using Tensorflow(Intermediate Level).

Task 4: Count Number of Objects Using OpenCV (Intermediate Level).

Task 5: Cat & Dog Classification Using CNN (Advanced level).

SCORES

Tasks	Level	Points
Task 1	Beginner	1
Task 2	Beginner	1
Task 3	Intermediate	4
Task 4	Intermediate	4
Task 5	Advanced	5

❖ SCORE 10/15 FOR SUCCESSFUL COMPLETION OF INTERNSHIP

TASK 1: Wine Quality Prediction

Objective

Predict wine quality using machine learning on a public dataset.

Task Breakdown

Setup Environment

Install Python and libraries: Pandas, Numpy, Seaborn, Matplotlib, Sklearn, XGBoost.

Model Development

1. Preprocess data (split, normalize).
2. Select models: Linear Regression, Decision Trees, Random Forest, XGBoost.
3. Train and evaluate models.

Model Optimization

1. Tune hyperparameters (Grid/Random Search).
2. Compare and select the best model.

Final Deployment

1. Train best model on entire dataset.
2. Save the trained model.

DOWNLOAD DATASET

TASK 2: Sales Forecasting

Objective

Predict daily sales using past sales data and other relevant inputs.

Task Breakdown

Data Preparation

1. Import libraries, Load and preprocess the dataset.
2. Transform date into 3 different inputs, Encode categorical features (season, festival).

Model Development

1. Split data into training and testing sets.
2. Define and compile a linear regression model using Keras.
3. Train the model on the training data.

Model Evaluation

Evaluate the model on the testing data, Visualize performance using matplotlib.

Prediction

1. Predict future sales using the trained model.
2. Multiply sigmoid output by maximum sales to get actual sales amount.

DOWNLOAD DATASET

TASK 3: Detecting Spam Emails Using Tensorflow

Objective

Build a TensorFlow-based model to classify emails as Spam or Ham.

Task Breakdown

Setup Environment

Install Python libraries: Pandas, Numpy, Matplotlib, Seaborn, Wordcloud, NLTK, TensorFlow.

Data Preparation

1. Import libraries, Load and preprocess dataset (clean, tokenize using NLTK).
2. Visualize data with Matplotlib, Seaborn, Wordcloud.

Model Development

1. Split data into training and testing sets.
2. Vectorize text using TF-IDF or word embeddings, Define and compile a TensorFlow model.

Model Training and Evaluation

Train the model, Evaluate performance (confusion matrix, accuracy).

Prediction

Classify new emails, Display results.

DOWNLOAD DATASET

TASK 4: Count Number of Objects Using Python-OpenCV

Objective

Count the number of objects in an image using OpenCV and Python.

Task Breakdown

Setup Environment

Install Python libraries: OpenCV, Numpy, Matplotlib.

Data Preparation

1. **Import Libraries:** OpenCV, Numpy, Matplotlib.
2. **Load Image:** Read the image using OpenCV.
3. **Preprocess Image:** Convert to grayscale, apply Gaussian blur.

Object Detection

1. **Thresholding:** Apply thresholding to create a binary image.
2. **Contour Detection:** Use OpenCV functions to find contours in the binary image.
3. **Count Objects:** Count the number of detected contours.

Visualization

1. **Draw Contours:** Highlight detected objects in the original image.
2. **Display Results:** Show the image with contours using Matplotlib.

DOWNLOAD IMAGE

TASK 5: Cat & Dog Classification Using CNN

Objective

Build a Convolutional Neural Network (CNN) to classify images as either cats or dogs.

Task Breakdown

Setup Environment

Install Python libraries: Pandas, Numpy, Matplotlib, Sklearn, OpenCV, TensorFlow.

Data Preparation

1. Pandas, Numpy, Matplotlib, Sklearn, OpenCV, TensorFlow, Load and preprocess cat and dog images.
2. Apply transformations (e.g., rotation, zoom) to increase dataset size.

Model Development

1. **Define Model:** Build a CNN using TensorFlow/Keras.
2. **Compile Model:** Set optimizer, loss function, and metrics.
3. **Train Model:** Train the CNN with training data.

Assess the model on validation/testing data, Plot training and validation accuracy/loss.

Prediction

1. **Classify Images:** Use the trained model to predict whether images are of cats or dogs.
2. **Display Results:** Show the predictions with corresponding images.

DOWNLOAD DATA

INSTRUCTION

❑You will need to complete 3 tasks (1 Beginner, 1 Intermediate and 1 Advanced level) and score 10/15 for successful completion of the internship.

❑Maintain a separate GitHub repository (name as “CodEvo Solutions ML” for all the tasks and share the link of the GitHub repo in the task submission form (it will be given later through email).

❑Update your LinkedIn profiles.

❑A TASK SUBMISSION FORM will be provided to you via email shortly. In the meantime, please proceed with your assigned tasks.

❑To showcase your work, create a video demonstrating your efforts and the outcomes of the tasks. The video should highlight your problem-solving skills, coding proficiency, and creativity in implementing solutions.

❑Once created, host the video on LinkedIn to showcase your work and build credibility among your peers and potential employers. Tag @CodEvoSolutions in your post and include the hashtag #CodEvoInternship to ensure maximum visibility.

❑Additionally, use relevant hashtags such as #MachineLearning, #Internship, and #CodingChallenge to reach a wider audience and attract potential collaborators or mentors.