<u>AI ASSSITED CODING</u>

<u>LABTEST-03</u>

NAME:M.SUCHITH

ENROLL.NO:2403A52068

BATCH.NO:04

SET E2:

QUESTION-01:

Scenario: In the domain of Agriculture, a company is facing a challenge related to data structures with ai.
Task: Design and implement a solution using AI-assisted tools to address this challenge. Include code, explanation of AI integration, and test results.
Deliverables: Source code, explanation, and output screenshots.

PROMPT:

You are an AI helping the farmer. Your task is to create a python script for manging sensor data for multiple fields detect anomalies and predict crop yield .use data structures efficiently and work on it.

CODE:

taask-01.py

labtest_03 > taask-01.py > ...

```python
from collections import deque, defaultdict
from datetime import datetime
import numpy as np
import random

"""
taask-01.py

Agriculture data-structures + AI example
- Purpose: manage sensor data for multiple fields, detect anomalies, and predict crop yield.
- Key data structures: CircularBuffer for recent sensor readings, DataRepository organizing per-field data.
- AI integration: lightweight linear regression (normal equation) implemented with numpy (no external ML libs).
  - Trains on aggregated sensor features to predict yield.
  - Simple statistical anomaly detector (z-score based) for incoming sensor readings.
- Test: synthetic dataset generation simulating sensor readings and yields; trains model and prints metrics.

How to run:
    python taask-01.py

Output (example):
    Trained model MSE: 3.12, R2: 0.88
    Test MSE: 2.95, R2: 0.90
    Example anomaly detection: flagged=True, z_scores=[...]
"""
# ---------------------
# Data structures
# ---------------------
class CircularBuffer:
    """Fixed-size FIFO buffer for recent sensor readings."""
    def __init__(self, size):
        self.size = size
        self.buf = deque(maxlen=size)

    def append(self, item):
        self.buf.append(item)

    def get_all(self):
        return list(self.buf)

```

labtest_03 > taask-01.py > ...

```python
class DataRepository:
    """Stores sensor readings per field and allows aggregation."""
    def __init__(self):
        # field_id -> list of (timestamp, reading_dict)
        self.store = defaultdict(list)

    def add_reading(self, field_id, reading):
        # reading: dict of feature_name -> float
        self.store[field_id].append((datetime.utcnow(), reading))

    def get_field_readings(self, field_id):
        return [r for (_, r) in self.store.get(field_id, [])]

    def aggregate_field_features(self, field_id, agg='mean'):
        # returns one aggregated feature vector per field (dict)
        readings = self.get_field_readings(field_id)
        if not readings:
            return {}
        keys = readings[0].keys()
        aggvals = {}
        for k in keys:
            vals = [r[k] for r in readings if k in r and r[k] is not None]
            if not vals:
                aggvals[k] = 0.0
            elif agg == 'mean':
                aggvals[k] = float(np.mean(vals))
            elif agg == 'median':
                aggvals[k] = float(np.median(vals))
            else:
                aggvals[k] = float(np.mean(vals))
        return aggvals

# ---------------------
# Lightweight AI model
# ---------------------

class LinearRegressor:
    """Linear regression via normal equation with optional regularization."""
    def __init__(self, l2=1e-6):
```

taask-01.py ●

labtest_03 > taask-01.py > ...

```python
class LinearRegressor:
    def __init__(self, l2=1e-6):
        self.theta = None
        self.l2 = l2
        self.feature_names = []

    def fit(self, X, y, feature_names=None):
        # X: (n_samples, n_features), y: (n_samples,)
        n, d = X.shape
        Xb = np.hstack([np.ones((n,1)), X])  # bias
        I = np.eye(d+1)
        I[0,0] = 0  # don't regularize bias
        self.theta = np.linalg.pinv(Xb.T.dot(Xb) + self.l2 * I).dot(Xb.T).dot(y)
        self.feature_names = feature_names or []
        return self

    def predict(self, X):
        n = X.shape[0]
        Xb = np.hstack([np.ones((n,1)), X])
        return Xb.dot(self.theta)

    def metrics(self, X, y):
        yhat = self.predict(X)
        mse = float(np.mean((y - yhat)**2))
        ss_res = float(np.sum((y - yhat)**2))
        ss_tot = float(np.sum((y - np.mean(y))**2)) if len(y) > 0 else 0.0
        r2 = 1 - ss_res/ss_tot if ss_tot > 0 else 0.0
        return {'mse': mse, 'r2': r2}

class AnomalyDetector:
    """Simple z-score based anomaly detector per feature."""
    def __init__(self, k=3.0):
        self.mean = None
        self.std = None
        self.k = k
        self.feature_names = []

    def fit(self, X, feature_names=None):
        # X: (n_samples, n_features)
```

labtest_03 > taask-01.py > ...

```python
class AnomalyDetector:
    def fit(self, X, feature_names=None):
        self.mean = np.mean(X, axis=0)
        self.std = np.std(X, axis=0, ddof=0) + 1e-9
        self.feature_names = feature_names or []
        return self

    def score(self, x):
        # x: (n_features,)
        z = np.abs((x - self.mean) / self.std)
        return z

    def is_anomaly(self, x):
        z = self.score(x)
        return bool(np.any(z > self.k)), z

# ---------------------
# Synthetic data generator
# ---------------------

def generate_synthetic_dataset(n_fields=50, samples_per_field=10, random_seed=42):
    random.seed(random_seed)
    np.random.seed(random_seed)
    repo = DataRepository()
    field_yields = {}
    # For each field, generate per-sample sensor readings and an overall yield
    for fid in range(n_fields):
        # true underlying coefficients per field (simulate variety)
        a_moist = random.uniform(0.5, 1.5)
        b_temp = random.uniform(-0.8, 0.2)
        c_rain = random.uniform(0.3, 1.0)
        d_fert = random.uniform(0.4, 1.2)
        intercept = random.uniform(5, 15)
        # generate samples
        soil_moistures = np.clip(np.random.normal(30 + 10*a_moist, 5, size=samples_per_field), 5, 60)
        temps = np.clip(np.random.normal(20 + 2*b_temp, 3, size=samples_per_field), -5, 45)
        humidity = np.clip(np.random.normal(60, 10, size=samples_per_field), 10, 100)
        rainfall = np.clip(np.random.exponential(2, size=samples_per_field), 0, 50)
        fertilizer = np.clip(np.random.normal(50 + 5*d_fert, 8, size=samples_per_field), 0, 200)
```

26 ⚠ 0 Ln 27, Col 25 Spaces: 2 UTF-8 CRLF {} Python 3.12.10 Go Li

labtest_03 > taask-01.py > ...

```python
def generate_synthetic_dataset(n_fields=50, samples_per_field=10, random_seed=42):

        # append readings to repo
        for i in range(samples_per_field):
            reading = {
                'soil_moisture': float(soil_moistures[i]),
                'temperature': float(temps[i]),
                'humidity': float(humidity[i]),
                'rainfall': float(rainfall[i]),
                'fertilizer': float(fertilizer[i])
            }
            repo.add_reading(fid, reading)

        # compute overall yield for the field as a linear combo + noise
        avg_m = np.mean(soil_moistures)
        avg_t = np.mean(temps)
        avg_r = np.mean(rainfall)
        avg_f = np.mean(fertilizer)
        noise = np.random.normal(0, 2.0)
        yield_val = intercept + a_moist*avg_m + b_temp*avg_t + c_rain*avg_r + d_fert*avg_f + noise
        field_yields[fid] = float(max(0.0, yield_val))  # yield non-negative

    return repo, field_yields

# ---------------------
# Utility: prepare dataset from repository
# ---------------------

def prepare_dataset(repo, field_yields, feature_keys=None):
    X_list = []
    y_list = []
    fkeys = feature_keys or ['soil_moisture','temperature','humidity','rainfall','fertilizer']
    for fid, y in field_yields.items():
        agg = repo.aggregate_field_features(fid, agg='mean')
        if not agg:
            continue
        row = [agg.get(k, 0.0) for k in fkeys]
        X_list.append(row)
        y_list.append(y)
```

26 ⚠ 0 Ln 27, Col 25 Spaces: 2 UTF-8 CRLF

labtest_03 > taask-01.py > ...

```python
def prepare_dataset(repo, field_yields, feature_keys=None):
    X = np.array(X_list, dtype=float)
    y = np.array(y_list, dtype=float)
    return X, y, fkeys

# ---------------------
# Tests / Example run
# ---------------------

def main():
    # generate synthetic data
    repo, yields = generate_synthetic_dataset(n_fields=80, samples_per_field=12)
    X, y, features = prepare_dataset(repo, yields)

    # split train/test
    n = X.shape[0]
    idx = np.arange(n)
    np.random.shuffle(idx)
    train_frac = 0.8
    cut = int(n * train_frac)
    train_idx, test_idx = idx[:cut], idx[cut:]
    X_train, y_train = X[train_idx], y[train_idx]
    X_test, y_test = X[test_idx], y[test_idx]

    # train regressor
    model = LinearRegressor(l2=1e-3).fit(X_train, y_train, feature_names=features)
    train_metrics = model.metrics(X_train, y_train)
    test_metrics = model.metrics(X_test, y_test)

    print("Trained model MSE: {:.3f}, R2: {:.3f}".format(train_metrics['mse'], train_metrics['r2']))
    print("Test MSE: {:.3f}, R2: {:.3f}".format(test_metrics['mse'], test_metrics['r2']))
    print("Learned parameters (bias + features):")
    # display bias and coefficients
    coef_names = ['bias'] + features
    for name, val in zip(coef_names, model.theta):
        print("  {:>12s}: {:.4f}".format(name, float(val)))

    # anomaly detector fit on training features
    ad = AnomalyDetector(k=3.0).fit(X_train, feature_names=features)
```

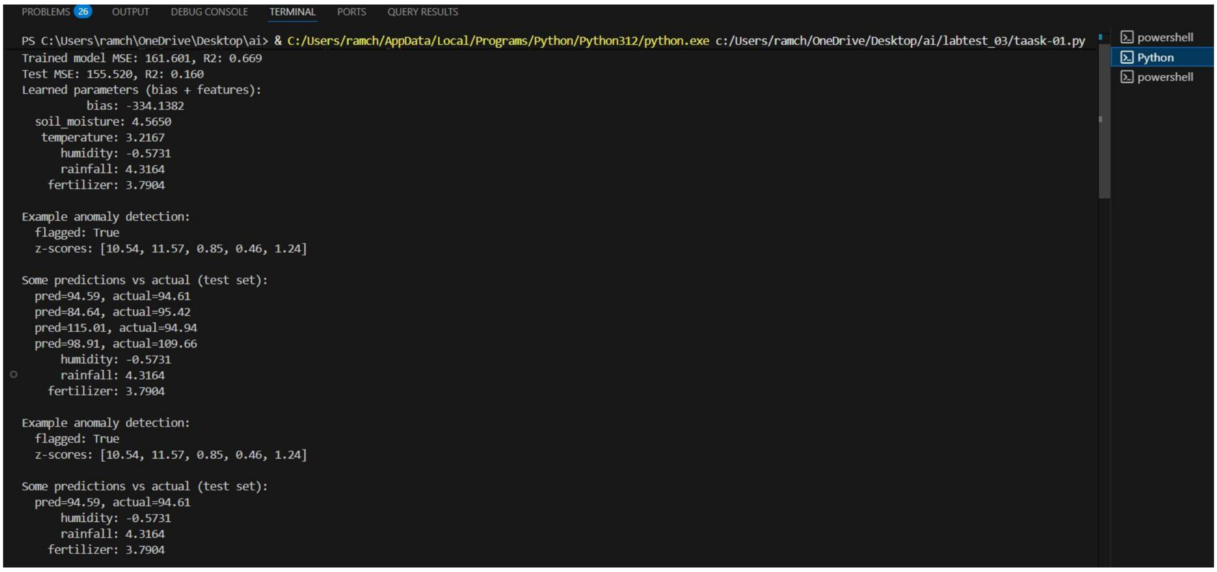
labtest_03 > taask-01.py > ...

```python
def main():
    # anomaly detector fit on training features
    ad = AnomalyDetector(k=3.0).fit(X_train, feature_names=features)
    # create an anomalous sample (extreme low moisture, high temp)
    anomalous_sample = X_test[0].copy()
    anomalous_sample[0] = anomalous_sample[0] * 0.1  # soil moisture dropped to 10%
    anomalous_sample[1] = anomalous_sample[1] + 15   # heat spike
    flagged, z_scores = ad.is_anomaly(anomalous_sample)
    print("\nExample anomaly detection:")
    print("  flagged:", flagged)
    print("  z-scores:", [round(float(z), 2) for z in z_scores])

    # show a few predictions vs actual
    print("\nSome predictions vs actual (test set):")
    preds = model.predict(X_test[:5])
    for i in range(min(5, len(preds))):
        print("  pred={:.2f}, actual={:.2f}".format(float(preds[i]), float(y_test[i])))

if __name__ == "__main__":
    main()
```

OUTPUT:



OBSERVATION:

The code given by the AI is efficient as it works properly. The AI had created an automated system which works on the managing sensor data for multiple fields ,detect anomalies and predict crop yield .

It predicts the data according to its sensation and it compares the predicted data with actual data.

The automated system acts as a fortune-teller for crops and it always spectates the crop as a watchdog for identifying the problems and helps to take the necessary steps to solve it.