AI-Assisted Farming for Crop Recommendation & Farm Yield Prediction Application

Team Venom

INTRODUCTION :

Overview

Agriculture is one of the oldest activities of civilization. Traditional farming is carried out based on knowledge passed from ancestors and personal experiences. Currently, the food production is inadequate and doesn't generate more income to farmers due to wrong crop cultivation according to farming factors like soil, climate, rainfall.

Purpose:

In this era of modern civilization and high population, the usage of machine learning tools will be useful for farmers and mankind.

LITERATURE SURVEY :

Existing Problem :

Several Recommendation Models are built according to Ensemble Learning Model using statistics and machine learning to achieve greater predictive efficiency.

Proposed Solution :

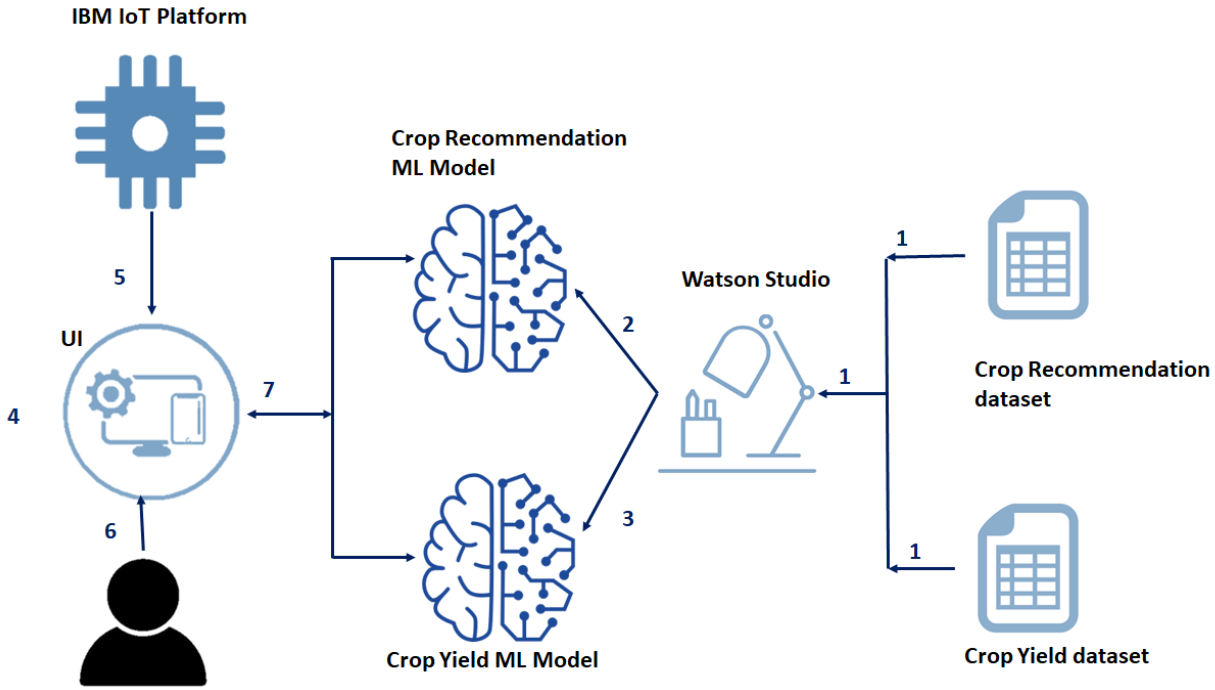
1. Cognos Analysis of Data using IBM Cognos Analytics.
2. Development Of Machine Learning Model using IBM WATSON STUDIO using AutoAI TOOL. Selection of algorithm

according to fastest build time and RMSE value.

3. Creation of NODE RED WEB APPLICATION and integration of machine learning model in the node red.
4. Development of a webpage and embedding NODE RED WEB APPLICATION in the webpage.
5. Development of Watson Assistant and integration in the webpage.

THEORITICAL ANALYSIS

BLOCK DIAGRAM



HARDWARE/SOFTWARE DESIGNING

HARDWARE :

The following table lists the minimum and recommended hardware requirements for the web application in personal computer

WEB APPLICATION HARDWARE REQUIREMENTS

Component	Minimum	Recommended
Processor	1.9 gigahertz (GHz) x86- or x64-bit dual core processor with SSE2 instruction set	3.3 gigahertz (GHz) or faster 64-bit dual core processor with SSE2 instruction set
Memory	2-GB RAM	4-GB RAM or more

Any Mobile Device can run the web application.

Network Requirements :

As Farming is at Remote Locations, Mobile Driven App is designed to work best over networks that have the following elements

- Bandwidth less than 50kbps
- Latency under 150 ms

SOFTWARE

Supported Web Browsers :

The web application can run in any of the following web browsers running on the specified operating systems of latest public release:

- Microsoft Edge
- Google Chrome
- Apple Safari

EXPERIMENTAL INVESTIGATIONS :
90% of data was used for training the algorithm and the rest 10% for testing purpose.

In case of crop production data, regression prediction type.

XGB Regressor algorithm generated RMSE 15920016.982,

Snap Decision Tree Regressor generated RMSE 168680035.189, 20708122.413, 21207842.478 and the fastest 8969848.870 with HPO-1, HPO-2, FE enhacement with 8 second of build time.

Model evaluation

Model evaluation measure

Measures	Holdout score	Cross validation score
	Root mean squared error	875746.452
R squared	0.074	0.573
Explained variance	0.075	0.574
Mean squared error	766931847913.926	98984122453868.078
Mean squared log error	18.702	15.619
Mean absolute error	121571.490	500136.616
Median absolute error	12556.901	13462.298
Root mean squared log error	4.325	3.901

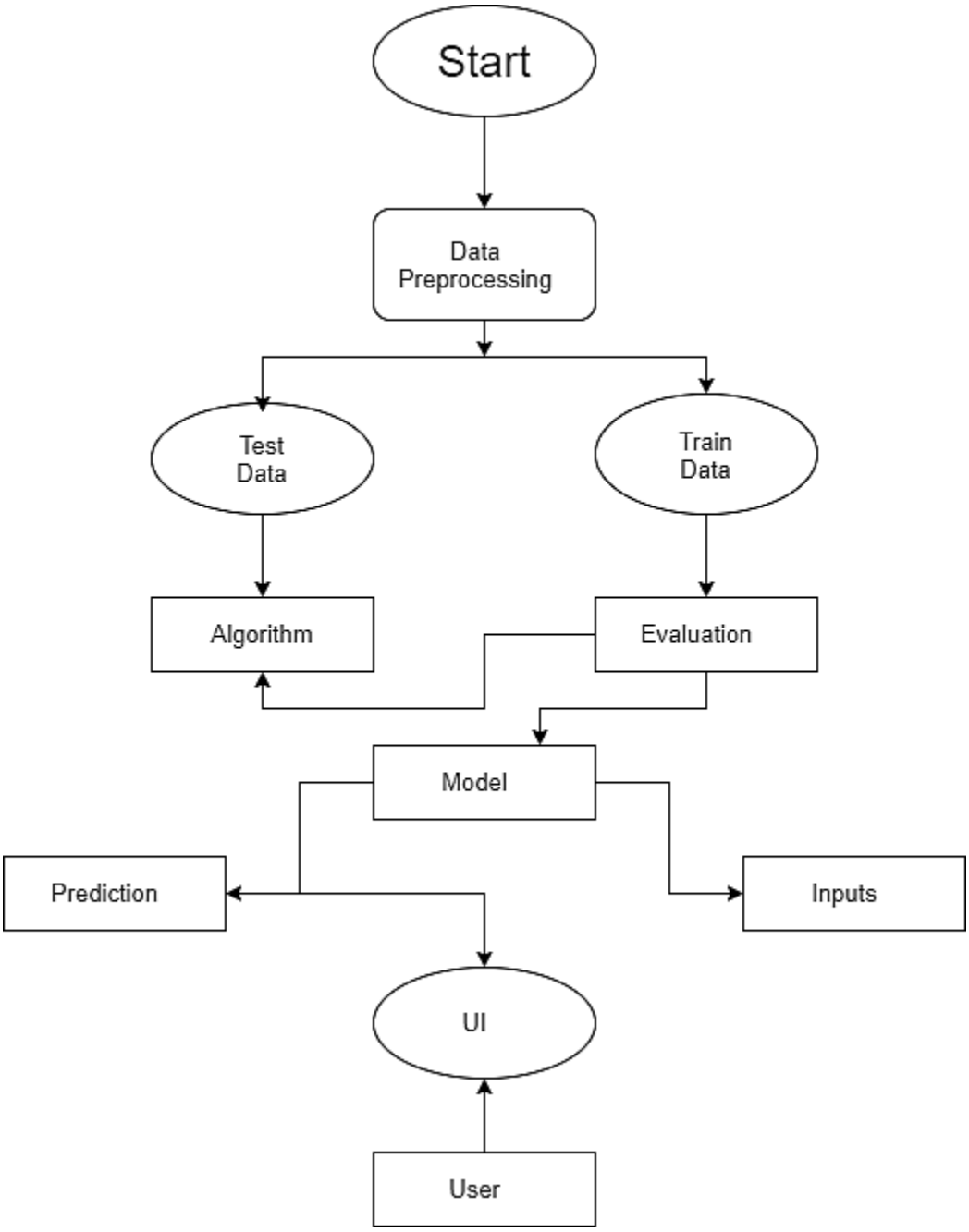
Multiclass Classification in case of crop recommendation data set.

Random Forest Classifier with HPO 1, FE, HPO-2 enhacements took 34

sec to build with 0.990 accuracy.

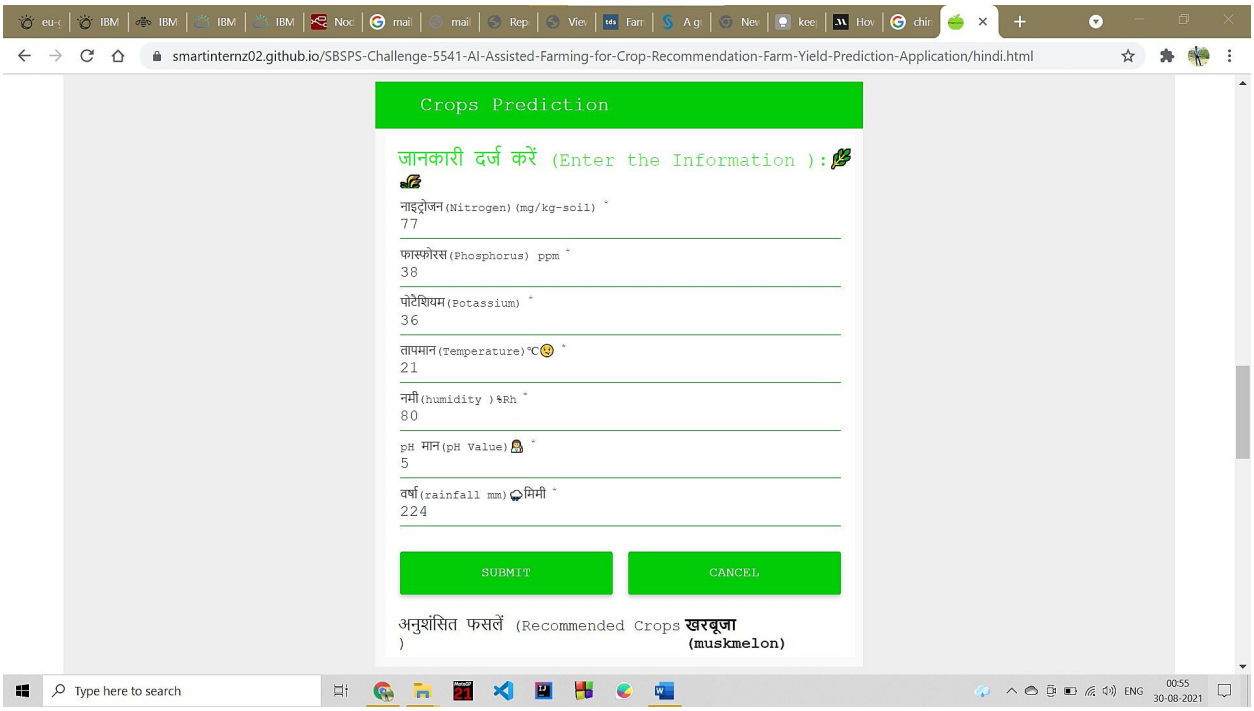
Snap Random Forest Classifier with 44 sec and 0.986 accuracy.

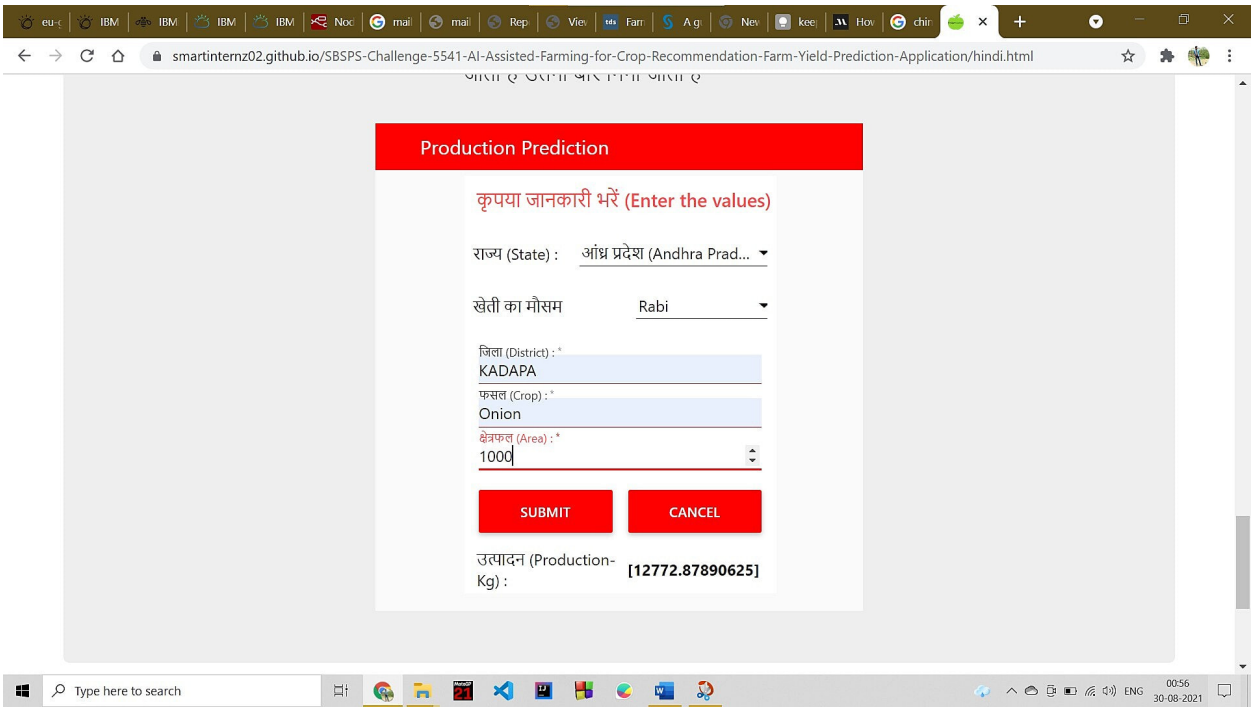
FLOWCHART :



RESULT :

Developed node red dashboards predicting crop and crop production. Also Deployed Watson Assistant for assistance on screen taps.





ADVANTAGES

Crop is predicted using Machine Learning Model which is highly accurate, similarly the total production prediction.

Data was analysed using regression and multiclass classification techniques which resulted in accurate predictions.

Integration of web application in a single webpage which is developed using mobile first approach leads to usage from any device.

Watson Assistant provides chat solution for user queries.

DISADVANTAGES

Limited Data leads to prediction inefficiency

Inconsistent Data also results into wrong prediction

Missing Values, Spelling Aberrations throughout the data are to be either removed or normalized to single value to improve model

accuracy

CUH LIMIT leads to limited utilization.

APPLICATIONS

This solution can be applied in the modern farming sector.

Accurate Crop and its production can be predicted.

CONCLUSION :

Machine Learning identifies trends and patterns in data reviewing large chunks of data and gives output in seconds.

Automation and continuous improvement are game changers in agricultural sector as we have learnt in IBM courses.

The capability to deliver much more in Agriculture sector is possessed by Machine Learning.

FUTURE SCOPE :

Application in Modern Farming Technique like plant species breeding and management, soil and water management, yield production and crop quality, disaster detection an weed detection.

BIBILOGRAPHY

https://www.mdpi.com/1424-8220/21/11/3758

https://farmer.gov.in/

APPENDIX

SOURCE CODE :

Package installation

```
!pip install ibm-watson-machine-learning | tail -n 1
!pip install -U autoai-libs==1.12.13 | tail -n 1
!pip install -U 'lale>=0.5.3,<0.6' | tail -n 1
!pip install -U scikit-learn==0.23.2 | tail -n 1
!pip install -U xgboost==1.3.3 | tail -n 1
!pip install -U lightgbm==3.1.1 | tail -n 1
!pip install -U snapml==1.7.4 | tail -n 1
```

Experiment metadata

```
from ibm_watson_machine_learning.helpers import DataConnection
from ibm_watson_machine_learning.helpers import S3Connection, S3Location

training_data_reference = [
    DataConnection(
    connection=S3Connection(
```

```
        api_key='pv9DKbYWpg5fZIWZ73u5N-1Njj-DgkBG9KYZj3_pTwZC',
        auth_endpoint='https://iam.bluemix.net/oidc/token/',
        endpoint_url='https://s3.eu.cloud-object-storage.appdomain.cloud'
    ),
        location=S3Location(
            bucket='crop-donotdelete-pr-zuzxlqhzorybla',
            path='Crop_recommendation.en.hi.csv'
        )
    ),
]
training_result_reference = DataConnection(
    connection=S3Connection(
        api_key='pv9DKbYWpg5fZIWZ73u5N-1Njj-DgkBG9KYZj3_pTwZC',
        auth_endpoint='https://iam.bluemix.net/oidc/token/',
        endpoint_url='https://s3.eu.cloud-object-storage.appdomain.cloud'
    ),
    location=S3Location(
        bucket='crop-donotdelete-pr-zuzxlqhzorybla',

path='auto_ml/3062baf4-0568-4434-a88e-afdea8abf7e5/wml_data/93
86ea63-d3d1-494f-a297-bf9e34738720/data/automl',

model_location='auto_ml/3062baf4-0568-4434-a88e-afdea8abf7e5/w
```

```
ml_data/9386ea63-d3d1-494f-a297-bf9e34738720/data/automl/pre_hpo_d_output/Pipeline1/model.pickle',

training_status='auto_ml/3062baf4-0568-4434-a88e-afdea8abf7e5/wml_data/9386ea63-d3d1-494f-a297-bf9e34738720/training-status.json'
    )
)


experiment_metadata = dict(
    prediction_type='multiclass',
    prediction_column='crop',
    holdout_size=0.1,
    scoring='accuracy',
    csv_separator=',',
    random_state=33,
    max_number_of_estimators=2,
    training_data_reference=training_data_reference,
    training_result_reference=training_result_reference,
    include_only_estimators=['RandomForestClassifier',
'DecisionTreeClassifier', 'LogisticRegression', 'ExtraTreesClassifier',
'GradientBoostingClassifier', 'XGBClassifier', 'LGBMClassifier',
'SnapDecisionTreeClassifier', 'SnapRandomForestClassifier',
'SnapLogisticRegression', 'SnapSVMClassifier'],
    deployment_url='https://eu-gb.ml.cloud.ibm.com',
```

```
    project_id='704cee98-5360-480f-8284-708db2f67056',
    drop_duplicates=True
)
```

Working with the completed AutoAI experiment

```
from ibm_watson_machine_learning.experiment import AutoAI

pipeline_optimizer = AutoAI(wml_credentials,
project_id=experiment_metadata['project_id']).runs.get_optimizer(met
adata=experiment_metadata)
```

Pipelines comparison

```
summary = pipeline_optimizer.summary()
best_pipeline_name = list(summary.index)[0]
summary
```

Deployment creation

```
target_space_id = "PUT_YOUR_TARGET_SPACE_ID_HERE"

from ibm_watson_machine_learning.deployment import WebService

service = WebService(
```

```
    source_wml_credentials=wml_credentials,
    target_wml_credentials=wml_credentials,
    source_project_id=experiment_metadata['project_id'],
    target_space_id=target_space_id
)
service.create(
    model=best_pipeline_name,
    metadata=experiment_metadata,
    deployment_name='Best_pipeline_webservice'
)
```

Running AutoAI experiment with Python AP

If you want to run the AutoAI experiment using the Python API, follow these. The experiment settings were generated basing on parameters set in the AutoAI UI.

- Go to your COS dashboard.
- In Service credentials tab, click New Credential.
- Add the inline configuration parameter: `{"HMAC":true}`, click Add. This configuration parameter adds the following section to the instance credentials, (for use later in this notebook):
- `"cos_hmac_keys": {`
- `"access_key_id": "***",`
- `"secret_access_key": "***"`
  `}`

Action: Please provide cos credentials in following cells.

- Use provided markdown cells to run code.

```
from ibm_watson_machine_learning.experiment import AutoAI

experiment = AutoAI(wml_credentials,
project_id=experiment_metadata['project_id'])
#@hidden_cell
cos_hmac_keys = {
    "access_key_id": "PLACE_YOUR_ACCESS_KEY_ID_HERE",
```

```
    "secret_access_key": "PLACE_YOUR_SECRET_ACCESS_KEY_HERE"
  }

cos_api_key = "PLACE_YOUR_API_KEY_HERE"
OPTIMIZER_NAME = 'custom_name'
from ibm_watson_machine_learning.helpers import DataConnection
from ibm_watson_machine_learning.helpers import S3Connection,
S3Location

training_data_reference = [
    DataConnection(
    connection=S3Connection(
        api_key='pv9DKbYWpg5fZIWZ73u5N-1Njj-DgkBG9KYZj3_pTwZC',
        auth_endpoint='https://iam.bluemix.net/oidc/token/',

endpoint_url='https://s3.eu.cloud-object-storage.appdomain.cloud',
        access_key_id = cos_hmac_keys['access_key_id'],
        secret_access_key = cos_hmac_keys['secret_access_key']
    ),
        location=S3Location(
            bucket='crop-donotdelete-pr-zuzxlqhzorybla',
            path='Crop_recommendation.en.hi.csv'
        )
    ),
]
training_result_reference = DataConnection(
    connection=S3Connection(
        api_key=cos_api_key,
        auth_endpoint='https://iam.bluemix.net/oidc/token/',

endpoint_url='https://s3.eu.cloud-object-storage.appdomain.cloud',
        access_key_id = cos_hmac_keys['access_key_id'],
        secret_access_key = cos_hmac_keys['secret_access_key']
    ),
    location=S3Location(
        bucket='crop-donotdelete-pr-zuzxlqhzorybla',

path='auto_ml/3062baf4-0568-4434-a88e-afdea8abf7e5/wml_data/9386ea
63-d3d1-494f-a297-bf9e34738720/data/automl',

model_location='auto_ml/3062baf4-0568-4434-a88e-afdea8abf7e5/wml_d
ata/9386ea63-d3d1-494f-a297-bf9e34738720/data/automl/pre_hpo_d_out
put/Pipeline1/model.pickle',
```

```
    training_status='auto_ml/3062baf4-0568-4434-a88e-afdea8abf7e5/wml_
    data/9386ea63-d3d1-494f-a297-bf9e34738720/training-status.json'
        )
    )
```

The new pipeline optimizer will be created and training will be triggered.

```
    pipeline_optimizer = experiment.optimizer(
        name=OPTIMIZER_NAME,
        prediction_type=experiment_metadata['prediction_type'],
        prediction_column=experiment_metadata['prediction_column'],
        scoring=experiment_metadata['scoring'],
        holdout_size=experiment_metadata['holdout_size'],
        csv_separator=experiment_metadata['csv_separator'],
        drop_duplicates=experiment_metadata['drop_duplicates'],

    include_only_estimators=experiment_metadata['include_only_estimato
    rs']
    )
    pipeline_optimizer.fit(
        training_data_reference=training_data_reference,
        training_results_reference=training_result_reference,
        background_mode=False,
    )
```