

SHRI VAISHNAV VIDYAPEETH VISHWAVIDYALAYA

**SHRI VAISHNAV INSTITUTE OF INFORMATION
TECHNOLOGY**

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING



PROJECT FILE

**STOCK MARKET FORECASTING USING AZURE MACHINE
LEARNING**

SESSION: January – June 2024

Submitted by:

Name : Tanushree Dour

Enrollment no. : 21100BTAIMLM09380

Year/ Sem : 3rd Year/ 6th Sem

Branch: B.Tech CSE(AIML)

Section: M

Submitted to:

Mr. Abhishek Shrivastava

TANUSHREE DOUR (21100BTAIMLM09380)

STOCK MARKET FORECASTING USING AZURE MACHINE LEARNING

Create an Azure Machine Learning workspace

1. Sign into the Azure portal using your Microsoft credentials.

The screenshot shows the 'Azure Machine Learning' workspace creation page in the Azure portal. The page is titled 'Azure Machine Learning' and has a subtitle 'Create a machine learning workspace'. It features a navigation bar with tabs: 'Basics', 'Networking', 'Encryption', 'Identity', 'Tags', and 'Review + create'. The 'Basics' tab is selected. Under 'Resource details', there is a note: 'Every workspace must be assigned to an Azure subscription, which is where billing happens. You use resource groups like folders to organize and manage resources, including the workspace you're about to create. [Learn more about Azure resource groups](#)'. Below this, there are two dropdown menus: 'Subscription' (selected: 'MOC Subscription--lod49071680') and 'Resource group' (selected: '(New) stockPrice', with a 'Create new' link). Under 'Workspace details', there is a note: 'Configure your basic workspace settings like its storage connection, authentication, container, and more. [Learn more](#)'. Below this, there are six dropdown menus: 'Name' (selected: 'StockPricePrediction', with a green checkmark), 'Region' (selected: 'East US 2'), 'Storage account' (selected: '(new) stockpricepred1252967190', with a 'Create new' link), 'Key vault' (selected: '(new) stockpricepred3317828530', with a 'Create new' link), 'Application insights' (selected: '(new) stockpricepred6433511333', with a 'Create new' link), and 'Container registry' (selected: 'None', with a 'Create new' link). At the bottom, there is a 'Review + create' button, a '< Previous' button, and a 'Next: Networking' button.

2. Select + **Create a resource**, search for Machine Learning, and create a new **Azure Machine Learning** resource with an Azure Machine Learning plan. Use the following settings:

- **Subscription:** Your Azure subscription.
- **Resource group:** Create or select a resource group.
- **Workspace name:** Enter a unique name for your workspace.
- **Region:** Select the closest geographical region.
- **Storage account:** Note the default new storage account that will be created for your workspace.
- **Key vault:** Note the default new key vault that will be created for your workspace.
- **Application insights:** Note the default new application insights resource that will be created for your workspace.
- **Container registry:** None (one will be created automatically the first time you deploy a model to a container).

IMPLEMENTING DATA SCIENCE SOLUTION ON AZURE

Microsoft Azure

Search resources, services, and docs (G+ /)

User1-40333404@LODS...
LODS-PROD-MSLEARN-MCA (LO...)

Home > Azure Machine Learning >

Azure Machine Learning

Create a machine learning workspace

Basics Networking Encryption Identity Tags Review + create

Resource details

Every workspace must be assigned to an Azure subscription, which is where billing happens. You use resource groups like folders to organize and manage resources, including the workspace you're about to create. [Learn more about Azure resource groups](#)

Subscription * ⓘ MOC Subscription--lod49071680

Resource group * ⓘ (New) stockPrice
[Create new](#)

Workspace details

Configure your basic workspace settings like its storage connection, authentication, container, and more. [Learn more](#)

Name ⓘ StockPricePrediction ✓

Region ⓘ East US 2

Storage account * ⓘ (new) stockpricepred1252967190
[Create new](#)

Key vault ⓘ (new) stockpricepred3317828530
[Create new](#)

Application insights ⓘ (new) stockpricepred6433511333
[Create new](#)

Container registry ⓘ None
[Create new](#)

[Review + create](#) < Previous Next: Networking

3. Select **Review + create**, then select **Create**. Wait for your workspace to be created (it can take a few minutes), and then go to the deployed resource.

Microsoft Azure

Search resources, services, and docs (G+ /)

User1-40333404@LODS...
LODS-PROD-MSLEARN-MCA (LO...)

Home >

Microsoft.MachineLearningServices | Overview

Deployment

Search

Delete Cancel Redeploy Download Refresh

Overview

Inputs

Outputs

Template

✓ Your deployment is complete

Deployment name : Microsoft.MachineLearningServices Start time : 5/3/2024, 11:24:10 AM

Subscription : MOC Subscription--lod49071680 Correlation ID : fa8258b5-f1dc-48e2-9a2d-b4f634ccf7ff

Resource group : stockPrice

> Deployment details

Next steps

[Go to resource](#)

Cost management

Get notified to stay within your budget and prevent unexpected charges on your bill. [Set up cost alerts >](#)

Microsoft Defender for Cloud

Secure your apps and infrastructure

Microsoft Azure

Search resources, services, and docs (G+ /)

User1-40333404@LODS...
LODS-PROD-MSLEARN-MCA (LO...)

Home > Microsoft.MachineLearningServices | Overview >

StockPricePrediction

Azure Machine Learning workspace

Search

Download config.json Delete

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Events

Settings

Monitoring

Automation

Support + troubleshooting

Essentials

Resource group : [stockPrice](#)

Location : East US 2

Subscription : [MOC Subscription--lod49071680](#)

Storage : [stockpricepred1252967190](#)

Studio web URL : <https://ml.azure.com?tid=bb7ed293-2674-4aef-a74a-dbf340a8ab...>


Container Registry : ...

Key Vault : [stockpricepred3317828530](#)

Application Insights : [stockpricepred6433511333](#)

MLflow tracking URI : <azureml://eastus2.api.azureml.ms/mlflow/v1/subscriptions/4149...>

JSON View



Work with your models in Azure Machine Learning Studio

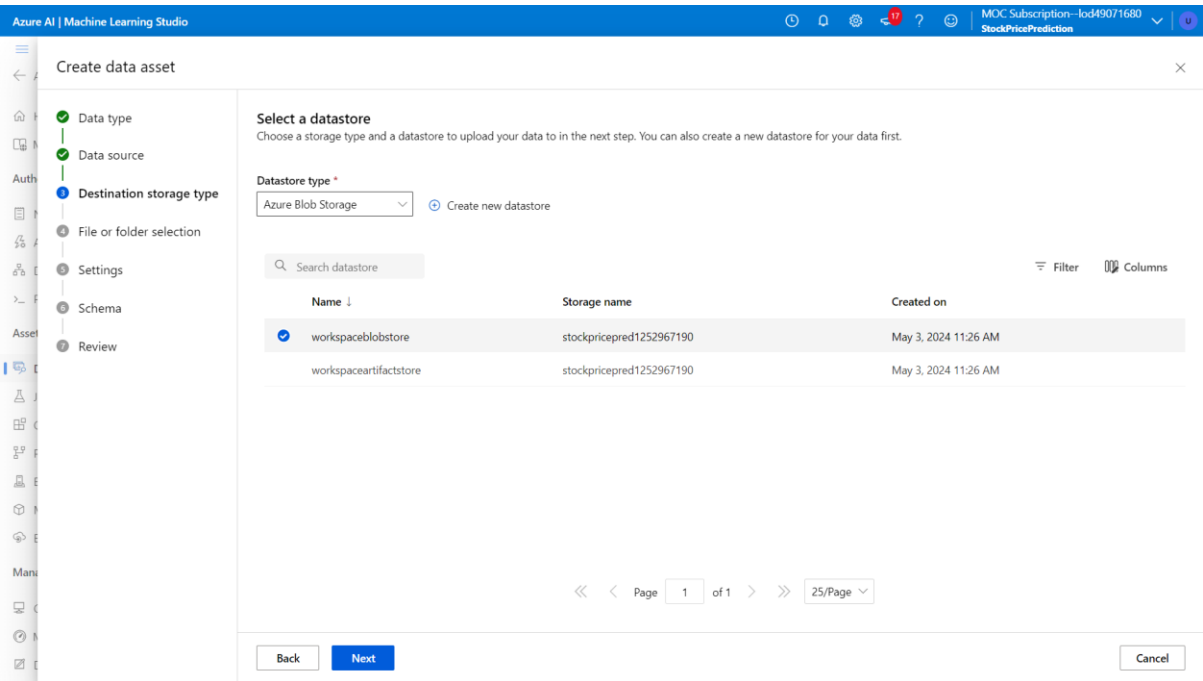
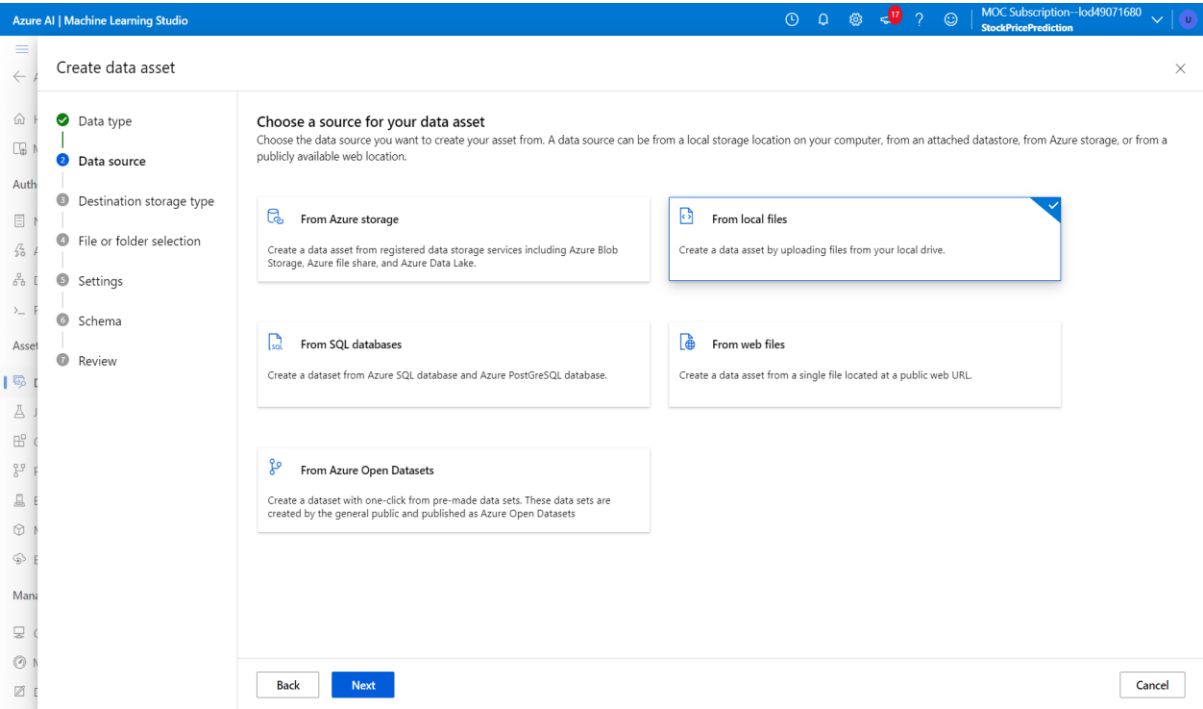
The Azure Machine Learning Studio is a web app where you can build, train, test, and deploy ML models. Launch it now to start exploring, or [learn more about the Azure Machine Learning studio](#)

[Launch studio](#)

TANUSHREE DOUR (21100BTAIMLM09380)

Create Compute

1. In [Azure Machine Learning studio](#), select the ≡ icon (a menu icon that looks like a stack of three lines) at the top left to view the various pages in the interface (you may need to maximize the size of your screen). You can use these pages in the left hand pane to manage the resources in your workspace. Select the **Compute** page (under **Manage**).



IMPLEMENTING DATA SCIENCE SOLUTION ON AZURE

Azure AI | Machine Learning Studio

⌚ 🔔 ⚙️ 17 ? 🗺️

MOC Subscription--lod49071680

StockPricePrediction

Create data asset

✓ Data type

✓ Data source

✓ Destination storage type

File or folder selection

Settings

Schema

Review

Choose a file or folder

Choose files or folders to upload from your local drive. If you upload multiple folders or files, they will be stored in a containing folder.

Upload path

azureml://subscriptions/4149bf5e-ab22-436e-8da5-87ac20679563/resourcegroups/stockPrice/works...

📁 Upload files or folder

☐ Overwrite if already exists

Upload list

Tesla_Stock_Prices.csv

161.41 KB/161.41 KB

...

Back Next Cancel

Information

What file types can I use?

Supported file types include: delimited (such as csv or tsv), Parquet, JSON Lines, and plain text.

Where are files uploaded?

Files will be uploaded to the selected datastore and made available in your workspace.

Azure AI | Machine Learning Studio

⌚ 🔔 ⚙️ 17 ? 🗺️

MOC Subscription--lod49071680

StockPricePrediction

Create data asset

✓ Data type

✓ Data source

✓ Destination storage type

✓ File or folder selection

Settings

Schema

Review

Settings

These settings determine how the data is parsed. The initial settings are automatically detected; you can change them as needed to reparse the data.

File format

Delimited

Delimiter

Comma

Example

Field1,Field2,Field3

Encoding

UTF-8

Column headers

All files have same headers

Skip rows

None

☐ Dataset contains multi-line data ⓘ

ⓘ Note: Processing tabular files with multi-line data is slower because multiple CPU cores cannot be used to ingest the data in parallel. Checking this option may result in slower processing times.

Data preview

Column1	Date	Open	High	Low	Close	Volume
0	2015-01-02 00:00:00	14.858	14.883	14.217	14.621	71466000
1	2015-01-05 00:00:00	14.303	14.433	13.811	14.006	80527500
2	2015-01-06 00:00:00	14.004	14.28	13.614	14.085	93928500
3	2015-01-07 00:00:00	14.223	14.319	13.985	14.063	44526000
4	2015-01-08 00:00:00	14.187	14.253	14.001	14.041	51637500
5	2015-01-09 00:00:00	13.928	13.999	13.664	13.777	70024500
6	2015-01-12 00:00:00	13.537	13.631	13.283	13.481	89254500
7	2015-01-13 00:00:00	13.555	13.841	13.394	13.617	67159500

Back Next Review Cancel

IMPLEMENTING DATA SCIENCE SOLUTION ON AZURE

Menu icons

Asset

Mani

Create data asset

✓ Data type

✓ Data source

✓ Destination storage type

✓ File or folder selection

✓ Settings

✓ Schema

Review

Review

Review the settings for your data asset and make any changes as needed.

Data type

Name

stock_price

Description

The dataset contains the stock prices of Tesla from 2015 to January 2024.

Type

tabular

Data source

Type

Local

File selection

Upload path

azureml://subscriptions/4149bf5e-ab22-436e-8da5-87ac20679563/resourcegroups/stockPrice/workspaces/StockPricePrediction/datastores/workspaceblobstore/paths/UI/2024-05-03_060226_UTC/Tesla_Stock_Prices.csv

Files uploaded

Tesla_Stock_Prices.csv

Storage

Datastore type

Schema

Column1	Integer
Date	Date
Open	Decimal
High	Decimal
Low	Decimal

(showing 5 of 8 columns)

Back

Create

Cancel

Azure AI | Machine Learning Studio

All workspaces

Home

Model catalog

Authoring

Notebooks

Automated ML

Designer

Prompt flow

Assets

Data

Jobs

Components

Pipelines

Environments

Models

Endpoints

Manage

Compute

Monitoring

Data Labeling

LODS-Prod-MSLearn-MCA > StockPricePrediction > Data > stock_price

stock_price

Version: 1 (latest)

Details Consume Explore Models Jobs

New version Refresh Generate profile Archive

Attributes

Type
Table (mitable)
Dataset type (from Azure ML v1 APIs)
Tabular
Created by
User1-40333404
Profile
[View profile](#)
Job: --
Files in dataset
1
Total size of files in dataset
161.4 KiB
Current version
1
Latest version
1
Created time
May 3, 2024 11:35 AM
Modified time
May 3, 2024 11:35 AM

Tags

No data

Description

The dataset contains the stock prices of Tesla from 2015 to January 2024.

Data sources

Datstore
[workspaceblobstore](#)

Relative path
UI/2024-05-03_060226.UTC/Tesla_Stock_Prices.csv

Actions
[View in datstores browse](#)
[View in Azure Portal](#)

Datstore URI
[azureml://subscriptions/4149bf5e-ab22-436e-8da5-87ac20679563/resourcegroups...](#)

Storage URI
[https://stockpricepred1252967190.blob.core.windows.net/azureml-blobstore-b9d6...](#)

TANUSHREE DOUR (21100BTAIMLM09380)

IMPLEMENTING DATA SCIENCE SOLUTION ON AZURE

Azure AI | Machine Learning Studio

LODS-Prod-MSLearn-MCA > StockPricePrediction > Data > stock_price

stock_price Version: 1 (latest) ☆

Details Consume Explore Models Jobs

Refresh Generate profile

Preview Profile

Number of columns: 7 Number of rows: 50 (of 2274)

Column1	Date	Open	High	Low	Close	Volume
0	2015-01-02 00:00:00	14.858	14.883	14.217	14.621	71466000
1	2015-01-05 00:00:00	14.303	14.433	13.811	14.006	80527500
2	2015-01-06 00:00:00	14.004	14.28	13.614	14.085	93928500
3	2015-01-07 00:00:00	14.223	14.319	13.985	14.063	44526000
4	2015-01-08 00:00:00	14.187	14.253	14.001	14.041	51637500
5	2015-01-09 00:00:00	13.928	13.999	13.664	13.777	70024500
6	2015-01-12 00:00:00	13.537	13.631	13.283	13.481	89254500
7	2015-01-13 00:00:00	13.555	13.841	13.394	13.617	67159500
8	2015-01-14 00:00:00	12.389	13.013	12.333	12.846	173278500
9	2015-01-15 00:00:00	12.966	13.05	12.667	12.791	78247500
10	2015-01-16 00:00:00	12.713	12.966	12.643	12.871	54048000
11	2015-01-20 00:00:00	12.925	12.941	12.469	12.795	67548000
12	2015-01-21 00:00:00	12.637	13.245	12.634	13.105	62295000
13	2015-01-22 00:00:00	13.133	13.549	13.013	13.441	61753500

Azure AI | Machine Learning Studio

LODS-Prod-MSLearn-MCA > StockPricePrediction > Training job

Submit an Automated ML job PREVIEW

Training method

Basic settings

Let's start with some basic information about your training job.

Job name *

Experiment name * ☐ ☒

New experiment name *

Description

Tags

Name : Value

Back Next

IMPLEMENTING DATA SCIENCE SOLUTION ON AZURE

Submit an Automated ML job

PREVIEW

Training method

Basic settings

Task type & data

Task settings

Compute

Review

Task type & data

Choose the type of task that you would like your model to perform and the data to use for training. [Learn more](#)

Select task type *
Regression

Select data
Make sure your data is preprocessed into a supported format.

Create

Refresh

Show supported data assets only

Reset view

Search

Filter

Columns

Name	Type	Created on	Modified on
stock_price	Table	May 3, 2024 11:35 AM	May 3, 2024 11:35 AM

Page 1 of 1 25/Page

Back

Next

Cancel

Submit an Automated ML job

PREVIEW

Training method

Basic settings

Task type & data

Task settings

Compute

Review

Task settings

Task type
Regression

Data
stock_price [View data](#)

Target column *
Close (Decimal)

View additional configuration settings

View featurization settings

Limits

Validate and test

You can choose a validation type and select test data as an optional step.

Validation type
Automatic

Test data
Train-test split

Percentage test of data *
10

Automated ML recommends that between 10 and 30 percent of data is held out for test

Back

Next

Cancel

IMPLEMENTING DATA SCIENCE SOLUTION ON AZURE

Virtual Machine

Advanced Settings

Standard_DS11_v2

Standard_DS3_v2

Standard_E4ds_v4

Standard_D13_v2

Select virtual machine

Select the virtual machine size you would like to use for your compute cluster.

Location *
East US 2

Virtual machine tier
☒ Dedicated ☐ Low priority

Virtual machine type
☒ CPU ☐ GPU

Virtual machine size
☒ Select from recommended options ☐ Select from all options

Name ↑	Category	Workload types	Available quota	Cost
Standard_DS11_v2 2 cores, 14GB RAM, 28GB storage	Memory optimized	Development on Notebooks (or other IDE) and light weight testing	300 cores	\$0.15/hr
Standard_DS3_v2 4 cores, 14GB RAM, 28GB storage	General purpose	Classical ML model training on small datasets	300 cores	\$0.23/hr
Standard_E4ds_v4 4 cores, 32GB RAM, 150GB storage	Memory optimized	Data manipulation and training on medium-sized datasets (1-10GB)	350 cores	\$0.29/hr
Standard_D13_v2 8 cores, 56GB RAM, 400GB storage	Memory optimized	Data manipulation and training on large datasets (>10 GB)	300 cores	\$0.60/hr

Back

Next

Virtual Machine

Advanced Settings

Standard_DS11_v2

Configure Settings

Configure compute cluster settings for your selected virtual machine size.

Name	Category	Cores	Available quota	RAM	Storage	Cost/Node
Standard_DS11_v2	Memory optimized	2	300 cores	14 GB	28 GB	\$0.15/hr

Compute name *
aml-cluster

Minimum number of nodes *
0

Maximum number of nodes *
1

Idle seconds before scale down *
120

☒ Enable SSH access

Advanced settings

Add tags
Name : Value

Add

☐ No tags

Back

Create

Download a template for automation.

Cancel

IMPLEMENTING DATA SCIENCE SOLUTION ON AZURE

Submit an Automated ML job

PREVIEW

Training method

Basic settings

Task type & data

Task settings

Compute

Review

Compute

Select and configure the compute resource for executing your training job.

Select compute type

Compute cluster

Select Azure ML compute cluster *

aml-cluster

+ New

Back

Next

Cancel

Submit an Automated ML job

PREVIEW

Training method

Basic settings

Task type & data

Task settings

Compute

Review

Review

Review or make changes to your job before submission.

Basic settings

Name

Stock_Market_Forecasting

Experiment name

Stock_Market_Forecasting

Description

--

Timeout (hours)

--

Tags

--

Task type & data

Task type

Regression

Data

stock_price

Compute settings

Compute type

Azure ML compute cluster

Selected Azure ML compute cluster

aml-cluster

Task settings

Target column

Close

Validate type

Automatic

Test data

Train-test split

Percentage test of data

10

Back

Submit training job

Cancel

IMPLEMENTING DATA SCIENCE SOLUTION ON AZURE

StockPricePrediction - Microso...

stock_price - Azure AI | Machin...

workspaceblobstore - Azure AI |

Stock_Market_Forecasting - Az...

Stock_Market_Forecasting - Az...

+

ml.azure.com/experiments/id/6337ef73-7881-49cd-8510-2e93c9536a5f/runs/stock_market_forecasting?wsid=.../subscriptions/4149bf5e-ab22-436e-8da5-87ac20679563/resourcegroups/...

MOC Subscription--lod49071680

StockPricePrediction

All workspaces

Home

Model catalog

Authoring

Notebooks

Automated ML

Designer

Prompt flow

Assets

Data

Jobs

Components

Pipelines

Environments

Models

Endpoints

Manage

86°F Sunny

LODS-Prod-MSLearn-MCA > StockPricePrediction > Jobs > Stock_Market_Forecasting > Stock_Market_Forecasting

Stock_Market_Forecasting Completed

Overview Data guardrails Models + child jobs Outputs + logs Child jobs

Refresh Edit and submit (preview) Register model Cancel Delete Compare (preview)

Status Completed

Warning: No scores improved over last 20 iterations, so experiment stopped early. This early stopping behavior can be disabled by

See more details

Created on May 3, 2024 11:43 AM

Start time May 3, 2024 11:43 AM

Duration 44m 2.159s

Compute duration 44m 2.159s

Compute target

Script name --

Created by User1-40333404

Job type Automated ML

Experiment Stock_Market_Forecasting

Arguments None

See all properties

Raw JSON

See YAML job definition

Inh YAML

Inputs

Input name: training_data
Data asset: stock_price:1
Asset URI: azureml:stock_price:1

Outputs

Output name: best_model
Model: azureml_stock_market_forecasting_40_output_mllflow_log_model_1010649396:1
Asset URI: azureml:azureml_stock_market_forecasting_40_output_mllflow_log_model_10...

Output name: full_training_dataset
Dataset: 5409a74e-0332-46bf-a496-8efaf2973bbc

Best model summary

Algorithm name VotingEnsemble

LODS-Prod-MSLearn-MCA > StockPricePrediction > Jobs > Stock_Market_Forecasting > Stock_Market_Forecasting

Stock_Market_Forecasting Completed

Overview Data guardrails Models + child jobs Outputs + logs Child jobs

Refresh Deploy Download Explain model View generated code Reset view

Search

Filter Columns

Algorithm name	Explained	Responsible AI	Normalized ro...	Sampling	Created on
VotingEnsemble			0.00730	100.00 %	May 3, 2024 12:26 PM
MinMaxScaler, ElasticNet			0.00754	100.00 %	May 3, 2024 11:51 AM
MaxAbsScaler, ElasticNet			0.00757	100.00 %	May 3, 2024 11:51 AM
StandardScalerWrapper, ElasticNet			0.00774	100.00 %	May 3, 2024 11:51 AM
MaxAbsScaler, ElasticNet			0.00799	100.00 %	May 3, 2024 11:51 AM
MaxAbsScaler, ElasticNet			0.00799	100.00 %	May 3, 2024 11:51 AM
RobustScaler, ElasticNet			0.00800	100.00 %	May 3, 2024 11:51 AM
StandardScalerWrapper, ElasticNet			0.00800	100.00 %	May 3, 2024 11:51 AM
StandardScalerWrapper, ElasticNet			0.00800	100.00 %	May 3, 2024 11:51 AM
MaxAbsScaler, ExtremeRandom Trees			0.00825	100.00 %	May 3, 2024 11:51 AM

Page 1 of 2 25/Page

IMPLEMENTING DATA SCIENCE SOLUTION ON AZURE

The screenshot displays the Azure Machine Learning Studio interface. The top navigation bar shows the path: StockPricePrediction > Jobs > Stock_Market_Forecasting > Stock_Market_Forecasting > lucid_clock_fb4pw6m4. The main panel shows the 'Model' tab for the job 'lucid_clock_fb4pw6m4', which is marked as 'Completed'. The 'Model summary' section displays the following details:

- Algorithm name: VotingEnsemble
- Ensemble details: [View ensemble details](#)
- Normalized root mean squared error: 0.00730 [View all other metrics](#)
- Sampling: 100.00 %
- Registered models: No registration yet
- Deployment status: No deployment yet

A 'Deploy a model' dialog box is open on the right, showing the configuration for deploying the model. The 'Name' field is set to 'stock-forecasting'. The 'Description' field is empty. The 'Compute type' is set to 'Azure Container Instance'. The 'Models' section lists 'stockmarketfore40'. The 'Enable authentication' checkbox is checked. The 'Use custom deployment assets' checkbox is unchecked. The 'Advanced' section is expanded, showing the following options:

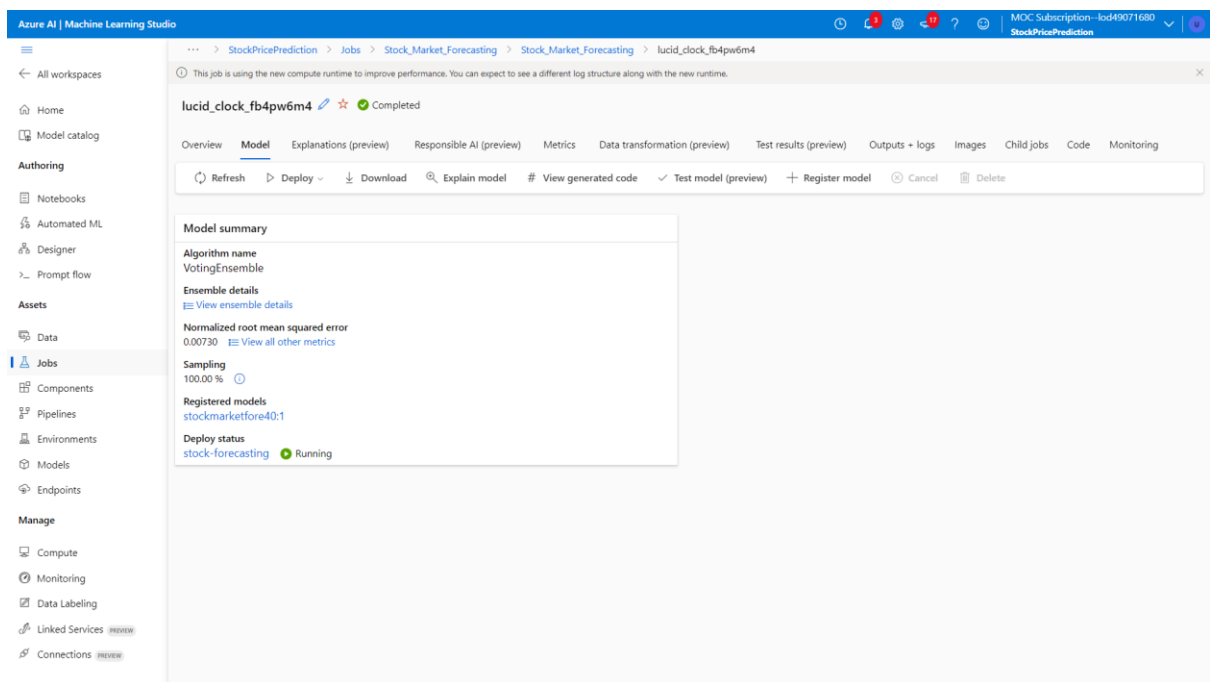
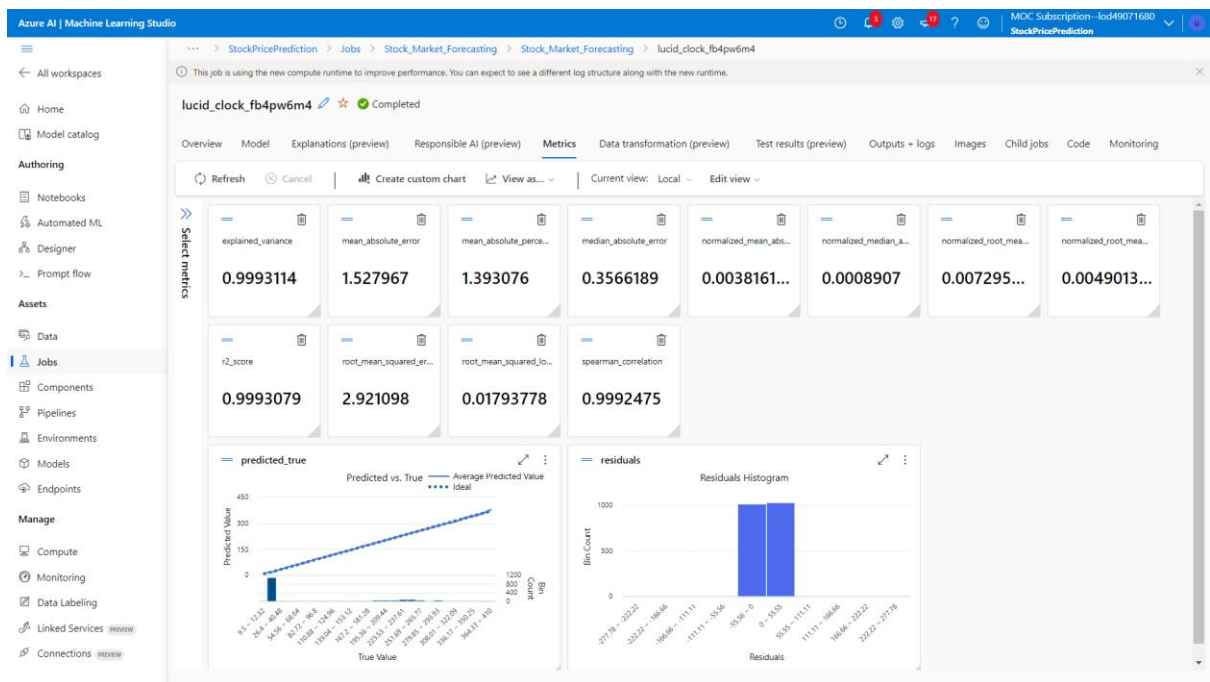
- Enable Application Insights diagnostics and data collection: ☐
- Enable Application Insights diagnostics and data collection: ☐
- Enable SSL: ☐
- Enable SSL: ☐
- CPU reserve capacity: 1.8
- Memory reserve capacity: 4

The bottom panel shows the 'Jobs' tab for the job 'lucid_clock_fb4pw6m4', which is marked as 'Completed'. The 'Model summary' section displays the following details:

- Algorithm name: VotingEnsemble
- Ensemble details: [View ensemble details](#)
- Normalized root mean squared error: 0.00730 [View all other metrics](#)
- Sampling: 100.00 %
- Registered models: [stockmarketfore40:1](#)
- Deployment status: [stock-forecasting](#) Running

TANUSHREE DOUR (21100BTAI MLM09380)

IMPLEMENTING DATA SCIENCE SOLUTION ON AZURE



TANUSHREE DOUR (21100BTAI MLM09380)

IMPLEMENTING DATA SCIENCE SOLUTION ON AZURE

