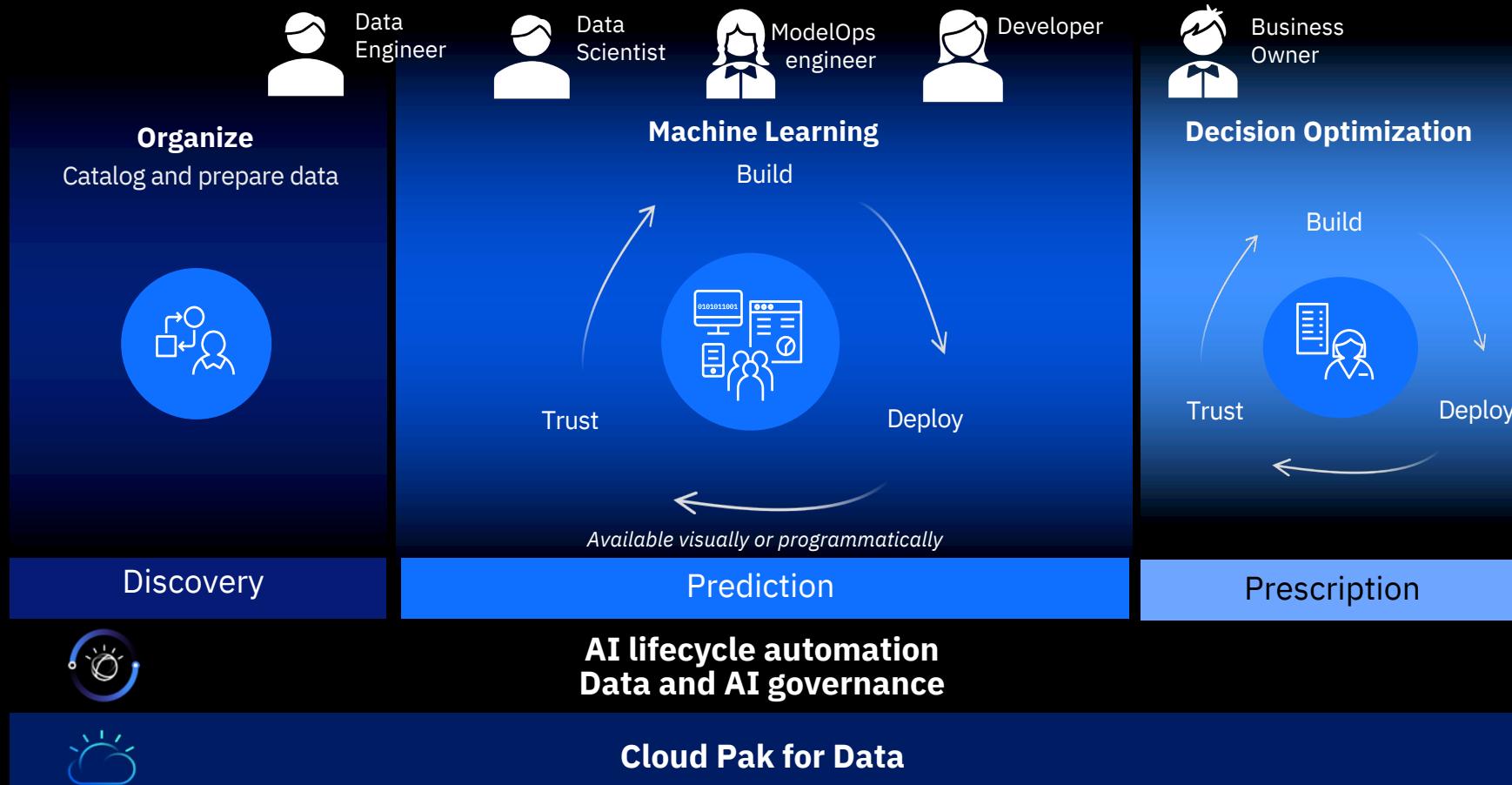


AI Life Cycle with Cloud Pak for Data – From Data to Predictions to Actions

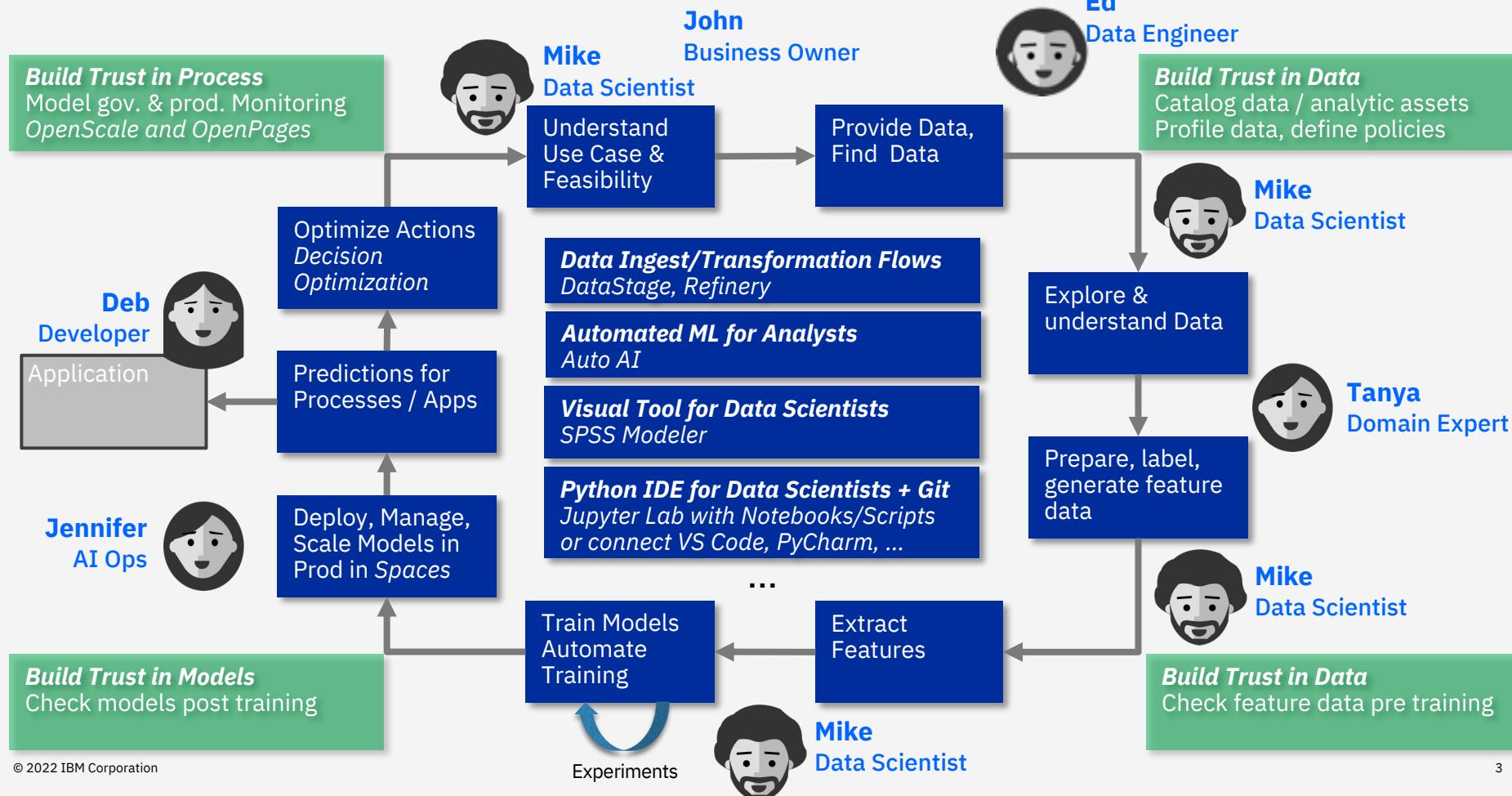
Thomas Schaeck

Distinguished Engineer, SAP Partnership and Watson Studio

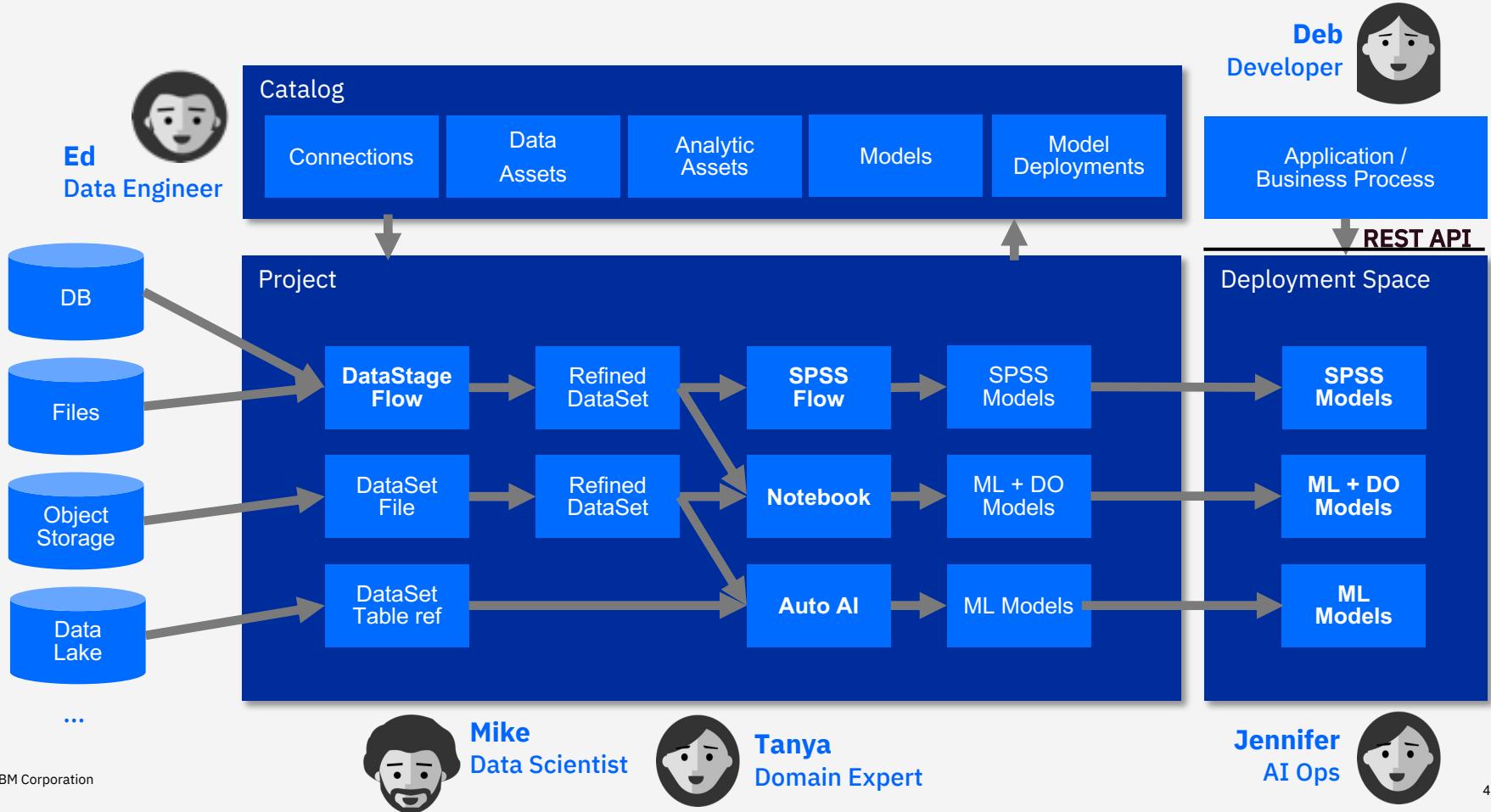
IBM Data Fabric – Trustworthy AI (Watson Studio)



Trusted AI Life Cycle with Data Fabric on Cloud Pak for Data



Collaborative Data Engineering, Data Science, and ML → Use in Apps



Projects with Git and IDEs



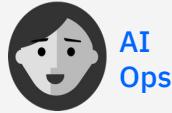
VS Code IDE +
WS Extension

Web Browser

VS Code
Py Charm
JupyterLab

Git-Based Project

Data Accessible
from Project



Space

VS Code Server

JupyterLab
Runtime Env.

RStudio
Runtime Env.

Git Repo

Code Files

CI/CD

Code
Package

Job UI

Job Asset JSON

Industrialize AI for Enterprises with automatic runtime/job management

Catalog(s)

Control who can use what data and other assets for what purpose

Projects

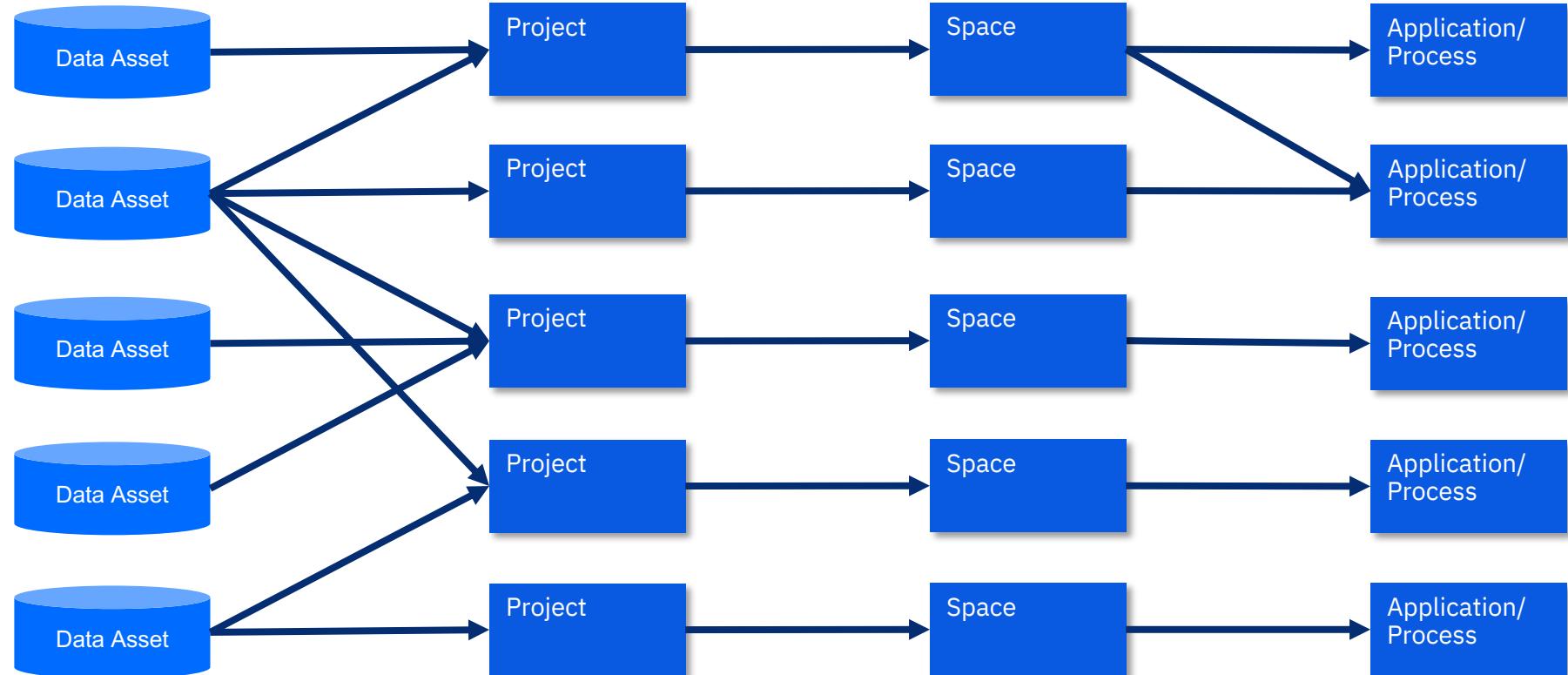
Analyze Data, use data to train ML models, optimize

Spaces

Deploy models for production use, monitor for fairness/drift

Solutions

Use models



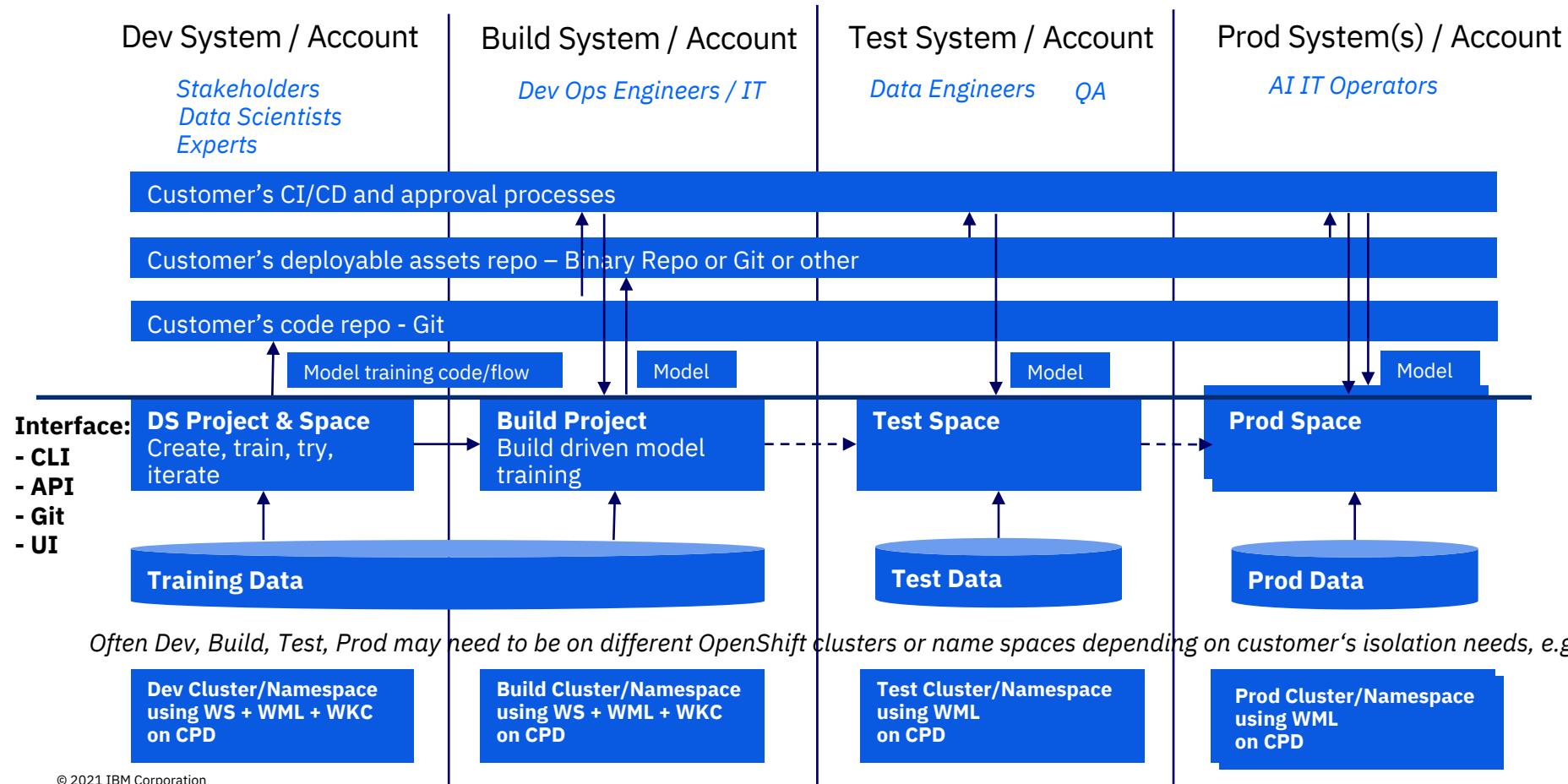
...

...

...

...

Example of Enterprise Deployment of CPD with CI/CD for Automation



Watson Studio on Cloud Pak for Data 4.6

New

Data Science & ML

Language & open source updates

- Python 3.10 (additional to Python 3.9)
- R4.2 (x86) / R 3.6 (Power)
- RStudio 2022.07.1

Enhanced VS Code integration connecting to WS Python runtimes on Cloud Pak for Data

Auto AI incremental training

supporting large training data 100GB+

Spark

- Spark 3.3 (deprecating 3.2)
- Job queueing when not enough free capacity
- Job prios to allow better controlling execution
- Spark live UI to view live state of spark jobs (additional to History server)

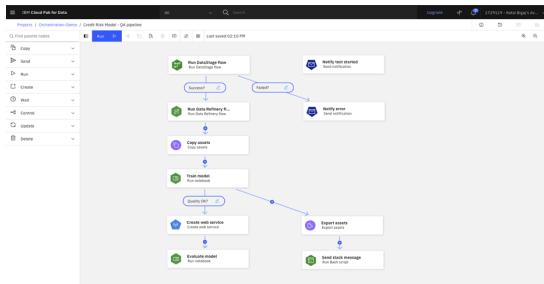
GA of Watson Natural Language Processing

libs for use in WS Python runtime environments

Watson Pipelines GA

Visually define pipelines to orchestrate

- Notebooks / Python Functions
 - Auto AI experiments
 - Refinery flows
 - DataStage flows
 - custom steps
- powered by Kubeflow pipelines



Watson Pipelines

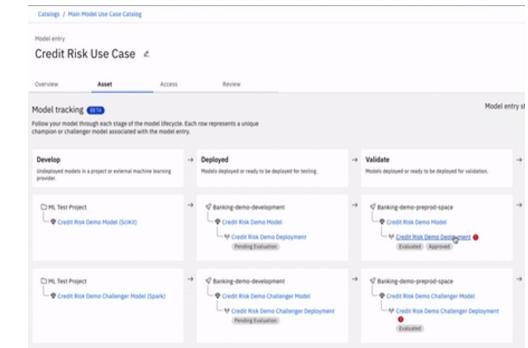
Responsible AI

Full OpenScale function for batch scoring in addition to online scoring

Model monitoring UI integration

Dashboards integrated with Space UI

Model Factsheets in installable on catalog or data science solution



Factsheets Model Repository

Welcome, Thomas!

Take a tutorial

Step through implementing a Data fabric use case in a sample project.

[Learn by example](#)

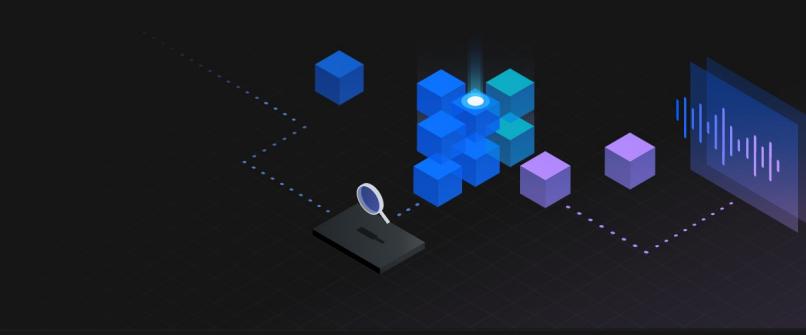
Work with data

Create a project for your team to prepare data, find insights, or build models.

[Create a project](#)

Extend your capabilities

Add tools, databases, or other features by creating services instances.

[Create a service](#)

Quick start

- Build dashboards with IBM Cognos Dashboard Embedded
- Create data pipelines with DataStage
- Build customer profiles with IBM Match 360 with Watson
- Catalog and govern data with Watson Knowledge Catalog
- Build and manage ML models with Watson Studio
- Query data anywhere with Watson Query

Overview

Projects		
	Demo Project	Today at 10:53 AM
	Flood risk analysis	Yesterday at 04:50 PM
	Train AutoAI and reference model - Demo	May 31, 2022 10:22 AM
	Credit Risk Fairness	May 20, 2022 11:32 AM

Catalogs		
	TEST - WCC	Jun 05, 2018 09:23 PM
	Product Team Watson Data Assets	Mar 24, 2018 05:56 PM
	New Test Catalog account schaeck	Oct 30, 2017 07:45 PM
	Great Outdoors Sandbox	Oct 30, 2017 07:11 AM
	Marketing Dept	Oct 13, 2017 02:46 AM

Notifications		
	Project import complete	Flood risk analysis was imported successfully. View import summary.
		Yesterday at 04:51 PM
	Asset promoted to space	Lukasz Cmielowski promoted credit risk fairness - P5 Snap Random Forest Classifier
		Yesterday at 03:44 PM
	Asset promoted to space	Rafal Bigaj promoted DB2-GCR_TRAINING_DATA_flow from project Tra
		Jun 02, 2022 08:43 AM

Support

Demo Catalog

Assets

Access control

Settings

Find assets

Filter by: Any asset type

Any tag

Clear all

Hide featured assets

Recently added

Watson recomme...

Highly rated



WML model

Credit Risk Model 1

Owner: Thomas Schaeck
Added: Jun 09, 2022 6:00 PM

Tags: cre...

☆☆☆☆☆ 0 reviews

Dashboard

Sales Dashboard

Owner: Thomas Schaeck
Added: Jun 09, 2022 5:55 PM

Tags:

★★★☆☆ 1 review

Notebook

German credit risk prediction wi...

Owner: Thomas Schaeck
Added: Jun 09, 2022 5:54 PM

Tags: cre...

★★★★★ 1 review

Data

German Credit Risk Training Data

Owner: Thomas Schaeck
Added: Jun 09, 2022 5:44 PM

Tags: cre...

★★★★☆ 1 review

Data

Titanic Data

Owner: Thomas Schaeck
Added: Jun 09, 2022 5:39 PM

Tags: titanic...

★★★★☆ 1 review

Showing 5 of 5 items

 Name

Owner

Tags

Business terms

Asset type

Date added

 Credit Risk Model 1 Thomas Schaeck

credit-risk

WML model

Jun 09, 2022

 German credit risk prediction with Scikit-learn Thomas Schaeck

credit-risk

Notebook

Jun 09, 2022

 German Credit Risk Training Data Thomas Schaeck

credit-risk

Data

Jun 09, 2022



Catalogs /

Demo Catalog



Data

German Credit Risk Training Data [Remove](#)[Add to project](#) [Overview](#)[Asset](#)[Access](#)[Review](#)[Profile](#)

Added: Jun 09, 2022 5:44 PM

Format: CSV [Tags](#) 

credit-risk

[Reviews](#) 1 review[Location:](#)IBM-Dallas [Sovereignty:](#) 

No sovereignty added yet.

[Connection](#)

Source: SAP Demo Catalog_DataCatalog

Source type: Cloud Object Storage

 GeneralDescription 

Data for training credit risk models.

Path 

No path

 Governance artifactsBusiness terms 

No business terms added yet.

Classifications 

No classifications added yet.

 Related assetsShow: Any relationship  Add asset 

Relationship	Asset name	Location	Asset type
Is related to	German credit risk prediction with Scikit-learn	Catalog \ Demo Catalog	Notebook
Is related to	Credit Risk Model 1	Catalog \ Demo Catalog	wml_model

Items per page: 25  1 of 1 pages

Catalogs

Add to project

Target*

Demo Project



Selected assets (1)

Asset name	Catalog	Connection
German Credit Risk Training Data	Demo Catalog	Demo Catalog_DataCatalog 

Cancel

Add

Catalogs / Demo Catalog

Data

German Credit Risk Training Data 

✓ 1 asset from Demo Catalog was added to project: Demo Project.

[Go to project](#)

Overview

Asset

Access

Review

Profile

Added: Jun 09, 2022 5:44 PM

Format: CSV Tags 

credit-risk

Reviews

 1 review

Location:

IBM-Dallas 

Sovereignty:

No sovereignty added yet. 

Connection

Source: Demo Catalog_DataCatalog

Source type: Cloud Object Storage

General

Description

Data for training credit risk models.



Path

No path

Governance artifacts

Business terms

No business terms added yet.



Classifications

No classifications added yet. 

Related assets

Show: Any relationship  Add asset 

Relationship	Asset name	Location	Asset type
Is related to	German credit risk prediction with Scikit-learn	Catalog \ Demo Catalog	Notebook
Is related to	Credit Risk Model 1	Catalog \ Demo Catalog	wml_model

Items per page: 25  1 of 1 pages

Projects

Find a project

New project



<input type="checkbox"/> Name	Date created	↓	Your role	Collaborators	
<input type="checkbox"/> Flood risk analysis	1 day ago		Admin		
<input type="checkbox"/> Fashion MNIST Project	1 month ago		Admin		
<input type="checkbox"/> 'Demo Project	2 months ago		Admin		
<input type="checkbox"/> Train AutoAI and reference model - Demo	2 months ago		Admin	 RB    +1	
<input type="checkbox"/> Data Exploration and Profiling	3 months ago		Admin	 AI   	
<input type="checkbox"/> Flood risk analysis T1	4 months ago		Admin		
<input type="checkbox"/> Flood risk analysis	4 months ago		Admin	 WS 	
<input type="checkbox"/> Unbox Demo	6 months ago		Editor	 SG  HP  +5	
<input type="checkbox"/> Predict customer interest to optimize a campaign with ML + DO 986987	6 months ago		Admin		
<input type="checkbox"/> Modeler AutoAI	7 months ago		Editor	 AL  MH  +2	
<input type="checkbox"/> Watson NLP Integration	8 months ago		Editor	 AL  SG  +2	
<input type="checkbox"/> Modeler FI	9 months ago		Admin	 AL 	
<input type="checkbox"/> Credit Card Fraud Project	10 months ago		Editor	 YD  HL  +5	

Items per page: 20 ▾ 1–20 of 83 items

1 ▾ 1 of 5 pages

Projects / Demo Project

Overview

Assets

Jobs

Manage



Launch IDE



Find assets

Add asset

New asset

13 assets

All assets

Name

Last modified

German Credit Risk Training Data CSV

4 minutes ago
Thomas Schaeck (You)

German credit risk prediction with Scikit-learn Notebook

17 minutes ago
Thomas Schaeck (You)

bank.csv CSV

7 hours ago



Titanic Sample Data application/x-ibm-rel-table

7 hours ago



Db2 Warehouse-i4 Connection

7 hours ago
Thomas Schaeck (You)

Python Notebook Notebook

2 months ago
Thomas Schaeck (You)

Orchestration Pipeline for Model Training Pipeline

2 months ago
Thomas Schaeck (You)

project_cos_connection Connection

2 months ago
Thomas Schaeck (You)

Bank marketing sample data - P4 Notebook Notebook

2 months ago
Thomas Schaeck (You)

Bank marketing sample data AutoAI experiment

2 months ago
Thomas Schaeck (You)

Decision Optimization Experiment CPLEX Decision Optimization experiment

2 months ago
Thomas Schaeck (You)

Sales Dashboard

2 months ago

Items per page: 20

1–13 of 13 items

1 of 1 pages



About this project

Name

Demo Project

**Description**

Demo

**Collaborators** Thomas Schaeck (You)
Admin**Controls**

Restricted eligibility

Cloud object storage

6.4 MB used

IBM Cloud accountName: Thomas Schäck's Account
ID: f8d190703020c2efd6f009d24c14d71c**SAML federation**

IBM w3id

Last modified

7 hours ago

Created on

Apr 01, 2022 by Thomas Schaeck

Find palette nodes



Save



Compile



Run



Loaded 6:34 PM



Connectors ^

Asset browser

Amazon RDS for MySQL

Amazon RDS for Oracle

Amazon RDS for PostgreSQL

Amazon Redshift

Amazon S3

Apache Cassandra

Apache Cassandra (optimized)

Apache HDFS

Apache Hive

Apache Kafka

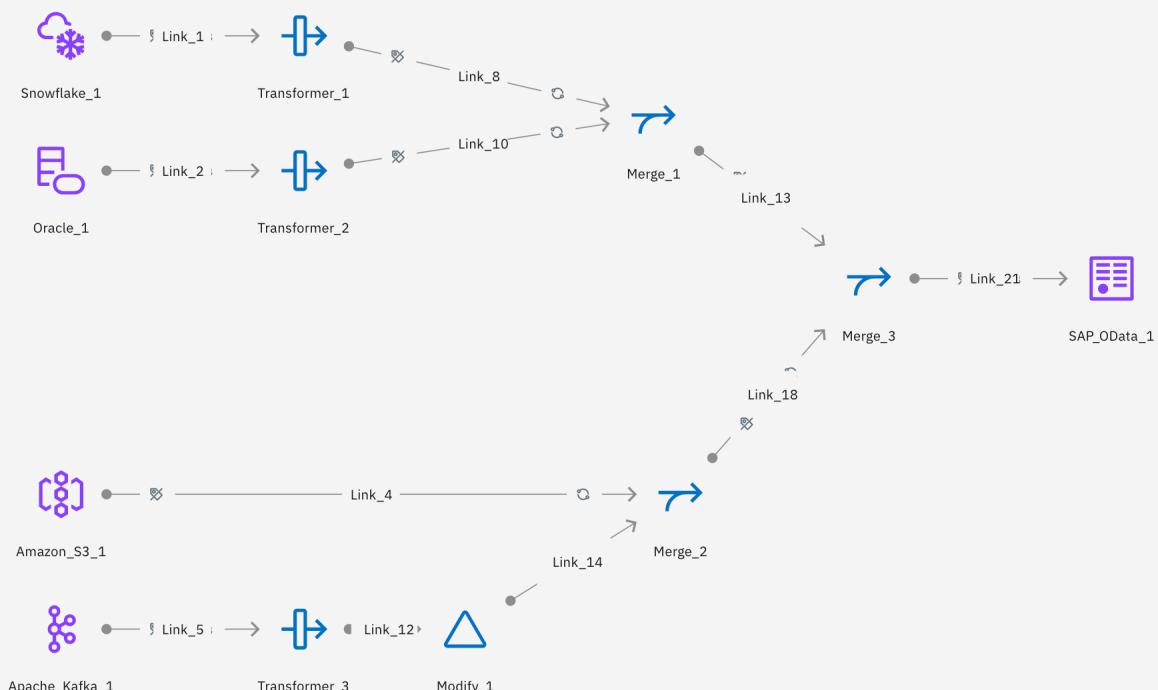
Box

Cloud Object Storage

Compose for MySQL

Data Virtualization

Data Virtualization



Preview

Profile

Visualization

Quality score 99% 165 classifiers

Current profile Last profile 9 Jun 2022 - 5:44 pm [View log](#)

Columns 21 Rows 5,000

Matches
Mismatch
Missing

[Delete profile](#)[Update profile](#)

CheckingStatus
Type: Varchar (1024)

Quality score 98%

• Code

100% 0% 0%

LoanDuration
Type: Varchar (1024)

Quality score 100%

• Code

100% 0% 0%

CreditHistory
Type: Varchar (1024)

Quality score 100%

• Text

100% 0% 0%

LoanPurpose
Type: Varchar (1024)

Quality score 98%

• Code

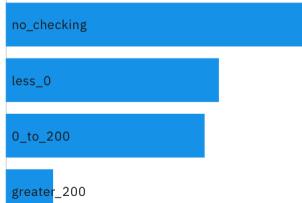
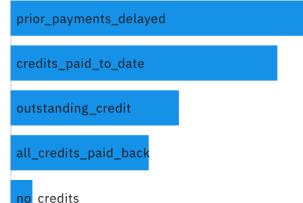
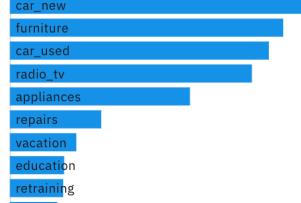
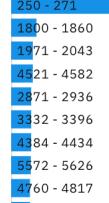
100% 0% 0%

LoanAmount
Type: Varchar (1024)

Quality score 99%

• Not classified

100%

Frequency**Frequency****Frequency****Frequency****Frequency****Statistics**

Unique 0
Minimum length 6
Maximum length 11

Statistics

Unique 6
Minimum length 1
Maximum length 2

Statistics

Unique 0
Minimum length 10
Maximum length 22

Statistics

Unique 0
Minimum length 5
Maximum length 10

Statistics

Unique 0
Minimum length 5
Maximum length 10

Information

Data asset

German Credit Risk Training Data

Description

Data for training credit risk models.

Tags

Creator

Thomas Schaeck

Usage

Created on Jun 09, 2022, 06:05 PM

Size

0

0

IBM Cloud Pak for Data

Search in your workspaces

Buy  1716985 - Thomas Schäck... 

Projects / Demo Project / Sales Promotion Study

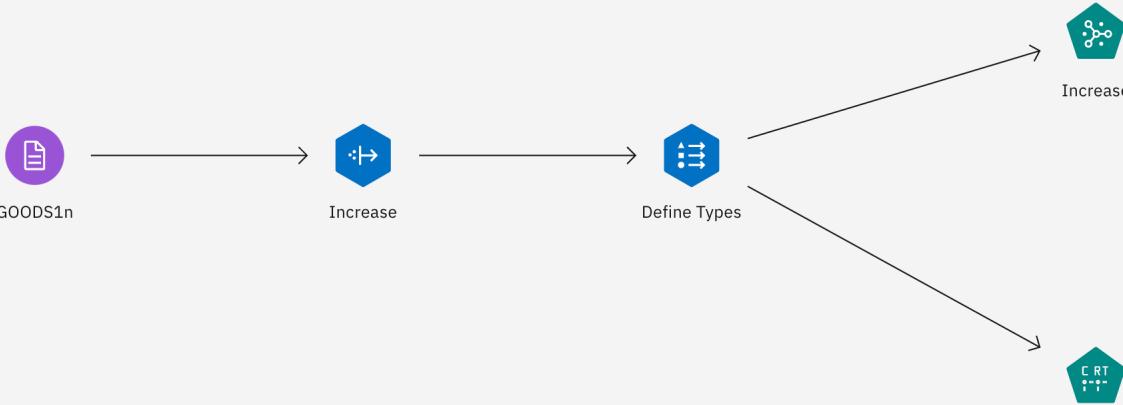
Find palette nodes Run selection 

Import Record Operations Field Operations Modeling Text Analytics Language Identifier Text Link Analysis Text Mining Graphs Charts Plot Multiplot Time plot Distribution Histogram Collection

GOODS1n  → Increase  → Define Types 

Increase 
Define Types 

Increase 
Increase 





Splitting the data into train and test

```
In [ ]: train_data, test_data = train_test_split(data_df, test_size=0.2)
```

Preparing the pipeline

```
In [ ]: features_idx = np.s_[0:-1]
all_records_idx = np.s_[:]
first_record_idx = np.s_[0]
```

In this step you will encode target column labels into numeric values. You can use `inverse_transform` to decode numeric predictions into labels.

```
In [ ]: string_fields = [type(fld) is str for fld in train_data.iloc[first_record_idx, features_idx]]
ct = ColumnTransformer([('ohe', OneHotEncoder(), list(np.array(train_data.columns)[features_idx][string_fields]))])
clf_linear = SGDClassifier(loss='log', penalty='l2', max_iter=1000, tol=1e-5)

pipeline_linear = Pipeline([('ct', ct), ('clf_linear', clf_linear)])
```

Train a model

```
In [ ]: risk_model = pipeline_linear.fit(train_data.drop('Risk', axis=1), train_data.Risk)
```

Evaluate the model

```
In [ ]: from sklearn.metrics import roc_auc_score

predictions = risk_model.predict(test_data.drop('Risk', axis=1))
indexed_preds = [0 if prediction=='No Risk' else 1 for prediction in predictions]

real_observations = test_data.Risk.replace('Risk', 1)
real_observations = real_observations.replace('No Risk', 0).values

auc = roc_auc_score(real_observations, indexed_preds)
```

Data

Files

Connections

Upload one file at a time. All file types accepted. 5 GB max file size.

Drag and drop files here or upload.

German Credit ...sk Training Data

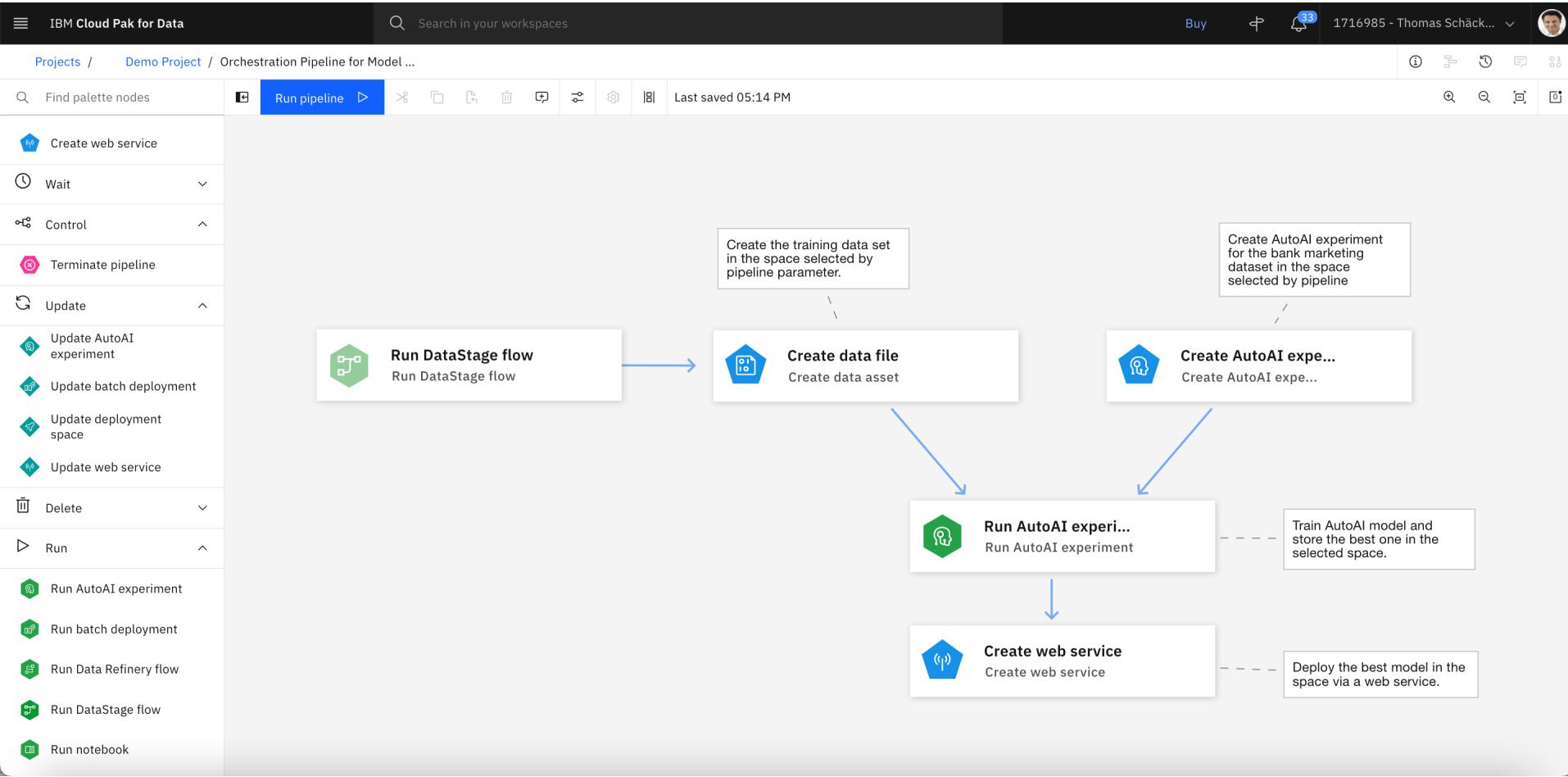
Insert to code 

pandas DataFrame

Credentials

bank.csv

Insert to code 



Auto AI

Preview

Profile

Visualization

Schema: 20 Columns

The preview includes only a limited set of columns and rows. Last refresh: 20 seconds ago 

Refine

CheckingS... String	LoanDura... String	CreditHis... String	LoanPurp... String	LoanAmo... String	ExistingSa... String	EmploymentDur... String	InstallmentPer... String	Sex String	OthersOn... String	CurrentResiden... String
0_to_200	12	none	none	none	100_to_500	none	3	female	none	none
0_to_200	1	none	none	none	100_to_500	none	3	female	none	none
0_to_200	3	none	none	none	100_to_500	none	3	female	none	none
0_to_200	2	none	none	none	100_to_500	none	3	female	none	none
0_to_200	4	none	none	none	100_to_500	none	3	female	none	none
0_to_200	5	none	none	none	100_to_500	none	3	female	none	none
0_to_200	6	none	none	none	100_to_500	none	3	female	none	none
0_to_200	7	none	none	none	100_to_500	none	3	female	none	none
0_to_200	8	none	none	none	100_to_500	none	3	female	none	none
0_to_200	9	none	none	none	100_to_500	none	3	female	none	none
0_to_200	10	none	none	none	100_to_500	none	3	female	none	none
0_to_200	11	none	none	none	100_to_500	none	3	female	none	none
0_to_200	31	credits_paid_to_l... other	other	1889	100_to_500	less_1	3	female	none	3
less_0	18	credits_paid_to_l... car_new	car_new	462	less_100	1_to_4	2	female	none	2

Information

 Data asset[german_credit_biased_no_ag...
e_unknown.csv](#)

Description

No description available for this asset

Tags

No tags available for this asset

Creator

Lukasz Cmielowski

Usage

Created on Jun 09, 2022, 01:27 PM

Size

680.871 KB

New asset

Select the tool to create an operational or configuration asset.

Tool type

All types

Automated builders

Graphical builders

Code editors

Component editors

Data access tools

Find tools by name or description

Automated builders



AutoAI

Automatically analyze your tabular data and generate candidate model pipelines customized for your predictive modeling problem.



Metadata enrichment

Enrich imported asset metadata with business context, data profiling, and quality assessment.

Graphical builders



Dashboard editor

Create a set of visualizations of analytical results on a graphical canvas without coding.



Data Refinery

Create a flow of ordered operations to cleanse and shape data. Visualize data to identify problems and discover insights.



Data privacy

Create and run masking flows to prepare copies of data assets that are masked by advanced data protection rules.



DataStage

Create a flow with a set of connectors and stages to transform and integrate data. Provide enriched and tailored information for your enterprise.



Decision Optimization

Create and manage scenarios to find the best solution to your optimization problem by comparing different



Pipelines

Automate the model lifecycle, including preparing data, training models, and creating deployments.



SPSS Modeler

Create a visual flow that uses modeling algorithms to prepare data and build and train a model, using a

Show descriptions ①

New asset

Create an AutoAI experiment

 New Gallery sample

Define details

Name

Credit Risk Prediction Experiment

Description (optional)

What's the purpose of this AutoAI experiment?

Tags (optional)

Add tags to make assets easier to find.

Start typing to add tags

Define configuration

Watson Machine Learning Service Instance

Machine Learning-vf

Environment definition 

8 vCPU and 32 GB RAM

This environment definition consumes **20 capacity units per hour** for training. The consumption rate differs for joining multiple data sources, depending on your configuration. For details, see [Watson Machine Learning plans](#).

Cancel

Create

Configure AutoAI experiment

Credit Risk Prediction Experiment

Add data sources



Add data sources

Drop or browse for one or more tabular data files. [Learn more.](#)

[Browse](#)

or

[Select from project](#)

Credit Risk Fairness data assets

Select up to 20 tabular data files. [Learn more](#)

Categories

2



Connection



Data asset



Data assets

3



cars4u_training.csv

 german_credit_biased_no_age_unkno... german_credit_data_biased_no_age.csv

Selected assets

1 / 1



All assets details must load before final selection.

german_credit_biased...



Asset name

german_credit_biased_no_age_unknown.cs
v

Asset details

Asset type: Data asset

Columns: 20

Size: 665 KB

Last modified:

2022/06/09 19:11:40

Created on:

2022/06/09 13:27:35

Cancel

Select asset

Configure AutoAI experiment

Credit Risk Prediction Experiment

Autosaved: 19:14:29

Add data sources

Drop or browse for one or more tabular data files. [Learn more.](#)

[Browse](#)

or

[Select from project](#) german_credit_biased_no_age_unknown.csv

Size: 0.68 MB | Columns: 20

Configure details



Create a time series forecast?

Enable this option to predict future activity over a specified date/time range. Data must be structured and sequential.

[Learn more](#) Yes No

What do you want to predict?

[Prediction columns !\[\]\(4a1237a65ad2b7a93414d17afab1c3c0_img.jpg\)](#)

Select prediction columns

INT ExistingCreditsCount

STR Job

INT Dependents

STR Telephone

STR ForeignWorker

STR Risk

Configure AutoAI experiment

Credit Risk Prediction Experiment 

Autosaved: 19:14:29

Add data sources

Drop or browse for one or more tabular data files. [Learn more.](#)[Browse](#)

or

[Select from project](#) german_credit_biased_no_age_unknown.csv

Size: 0.68 MB | Columns: 20

Configure details



Create a time series forecast?

Enable this option to predict future activity over a specified date/time range. Data must be structured and sequential.

[Learn more](#) Yes No

What do you want to predict?

Prediction columns 

Risk

Prediction column: Risk

CUH used this month: 187.66 CUH

PREDICTION TYPE

Multiclass Classification

OPTIMIZED FOR

Accuracy & run time

Experiment settings Run experiment 

Experiment settings

Prediction column
Risk (STR)Data source
german_credit_biased_no_age_unknown.csv

Prediction

Data source

Runtime

Prediction settings

General

Fairness

Time series

Fairness evaluation

Enable this option to check each pipeline for bias by calculating the disparate impact ratio. This tracks whether a pipeline shows a tendency to provide a favorable (preferred) outcome for one group more often than another.

 Enable fairness evaluation

Favorable outcomes

Specify the classes from your prediction column that would be considered favorable. Classes left unspecified will be assumed as unfavorable.

Enter favorable outcomes

Add class

Suggested classes are based on a subset of your data source. Not all classes may appear.

Favorable outcomes for [Risk]

No Risk

Protected attribute method

Choose how to evaluate features that are potential sources of bias.

Auto

Experimental

Manual

Cancel

Save settings

Experiment settings

Prediction column

Risk (STR)

Data source

german_credit_biased_no_age_unknown.csv

X

Prediction

General

Fairness

Time series

Auto Experimental

AutoAI will detect common protected attributes, such as race, gender, and age.

Manual

Manually specify protected attributes.



Protected attributes ⓘ

1 × Sex



Sex

male, female ⌂

Select a class from **Sex**, then indicate whether the class is a monitor or reference group.

Class

Group type

Select class

Reference

Add group +

Select class

Group

Group type

female

Monitored



male

Reference



Cancel

Save settings

IBM Cloud Pak for Data

Search in your workspaces

Buy 32 1716985 - Thomas Schäck... 

Projects / Credit Risk Fairness / Credit Risk Prediction Experiment

Experiment summary Pipeline comparison ★ Rank by: Accuracy (Optimized) | Cross validation score

Progress map ⓘ
Prediction column: Risk

Relationship map

Swap view ↗

Preprocessing
GERMAN_CREDIT_BIAS...
Setting default preprocessor parameters
Time elapsed: 92 seconds

[View log](#) [Save code](#)

Pipeline leaderboard ⚡

Rank	↑	Name	Algorithm	Accuracy (Optimized) Cross Validation	Enhancements	Build time
...

...

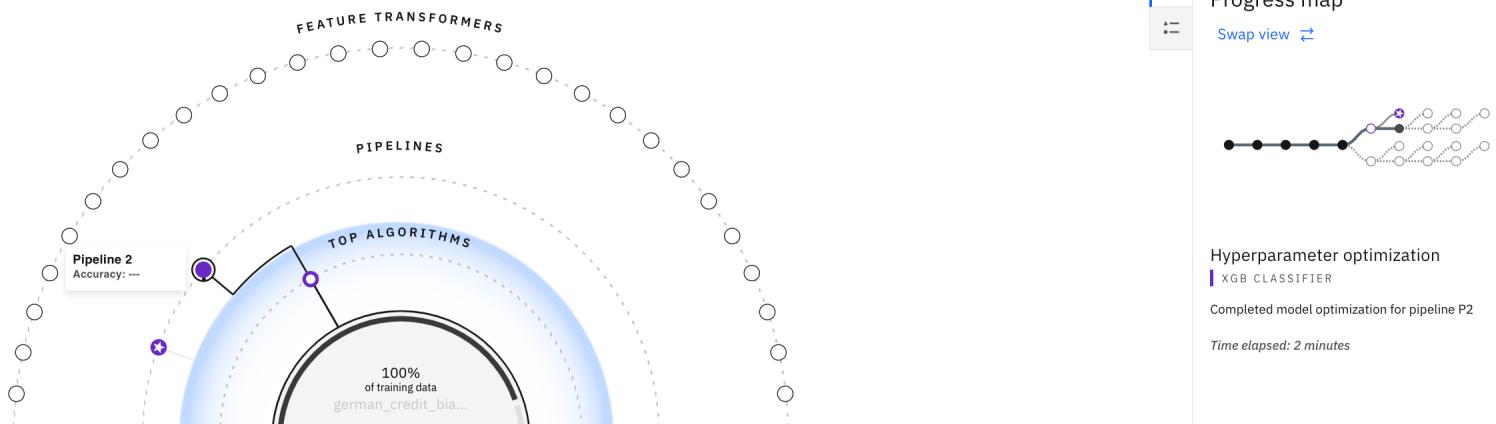
Experiment summary

Pipeline comparison

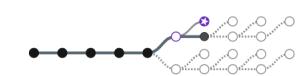
★ Rank by: Accuracy (Optimized) | Cross validation score

Relationship map 

Prediction column: Risk



Progress map

Swap view 

Hyperparameter optimization

XGB CLASSIFIER

Completed model optimization for pipeline P2

Time elapsed: 2 minutes

View log

Save code

Pipeline leaderboard 

Rank	↑	Name	Algorithm	Accuracy (Optimized) Cross Validation	Disparate impact Cross Validation	Enhancements	Build time
------	---	------	-----------	--	--------------------------------------	--------------	------------

 1

Pipeline 1

 XGB Classifier

0.789

1.019

None

00:00:03

Pipeline details

Pipeline 7

Rank

Accuracy and disparate impact (Optimized)

Algorithm

Enhancements

1 0.877 (Holdout)

XGB Classifier

HPO-1 FE

Save as

Model viewer

Model information

Feature summary

Evaluation

Model evaluation

Fairness

Confusion matrix

Precision recall

Threshold

Feature summary

High correlation

All features

Search feature or transformer names

Feature name

Transformation

Feature importance

OthersOnLoan	None	100.00%
ExistingCreditsCount	None	60.00%
Telephone	None	58.00%
LoanDuration	None	41.00%
CurrentResidenceDuration	None	39.00%
OwnsProperty	None	37.00%
CheckingStatus	None	32.00%
EmploymentDuration	None	28.00%
Sex	None	23.00%
InstallmentPlans	None	19.00%
Housing	None	18.00%

Pipeline details

Pipeline 7

Model viewer

Model information

Feature summary

Evaluation

Model evaluation

Fairness

Confusion matrix

Precision recall

Threshold

Rank

Accuracy and disparate impact (Optimized)

Algorithm

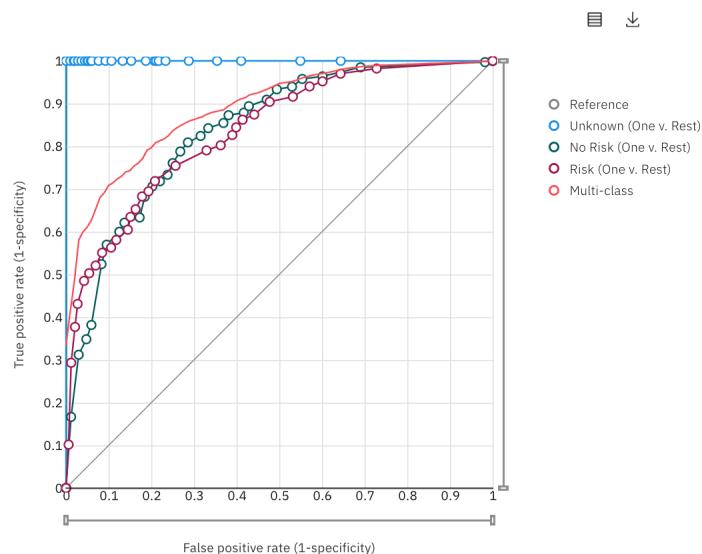
Enhancements

1 0.877 (Holdout)

XGB Classifier

HPO-1 FE

Save as

Model evaluation ROC curve 

Pipeline details

Pipeline 7

Rank

Accuracy and disparate impact (Optimized)

Algorithm

Enhancements

1

0.877 (Holdout)

XGB Classifier

HPO-1 FE

Save as

Model viewer

Model information

Feature summary

Evaluation

Model evaluation

Fairness

Confusion matrix

Precision recall

Threshold

Fairness

Overall fairness

Most fair

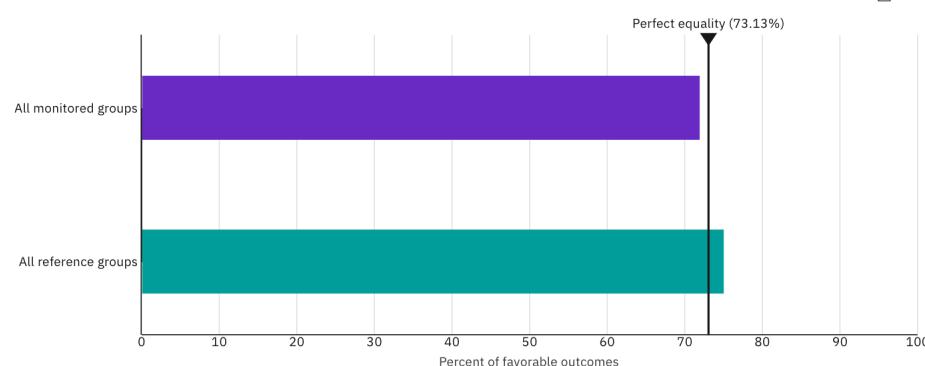
Favorable outcome

No Risk

Disparate impact ratio

Threshold : 80% | Actual: 95.85% (15.85%) [View calculation](#)

Fairness summary



Pipeline details

Pipeline 7

Rank

Accuracy and disparate impact (Optimized)

Algorithm

Enhancements

1

0.877 (Holdout)

XGB Classifier

HPO-1 FE

Save as

Model viewer

Model information

Feature summary

Evaluation

Model evaluation

Fairness

Confusion matrix

Precision recall

Threshold

Confusion matrix

View

Multi-class

Observed	Predicted			
	No Risk	Risk	Unknown	Percent correct
No Risk	297	69	0	81.1%
Risk	33	98	0	74.8%
Unknown	0	0	1	100.0%
Percent correct	90.0%	58.7%	100.0%	79.5%

Less correct

More correct

 Find assetsAdd asset New asset 

11 assets

 All assets

Asset types

>  Data 3<  Experiments 5

AutoAI experiment 5

>  Source Code 2<  Models 1

Model 1

Model

 Name

Type

Software specification

Last modified credit risk fairness - P5 Snap Random Forest Classifier
Model

wml-hybrid_0.1

hybrid_0.1

1 day ago
Lukasz CmielowskiDrop data files here or
browse for files to upload

Items per page:

20



1–1 of 1 items

1 of 1 pages



Deploy Model to Deployment Space

New deployment space

Use a space to collect assets in one place to create, run, and manage deployments

Define details

Name

Description (Optional)

Deployment space tags (optional)

Select storage service

Select machine learning service (optional)

Upload space assets (optional)

Populate your space with assets exported from a project or space to a .zip file. You can add more assets after the space is created.

Drop .zip file here or browse your files to upload

CancelCreate

Promote to space

Use a deployment space to organize supporting resources such as input data and environments; deploy models or functions to generate predictions or solutions; and view or edit deployment details.

Target space

Demo Deployment Space

Tags (optional)

Start typing to add tags



Why don't I see all of my spaces?



Go to the model in the space after promoting it

Selected assets (1)

Asset name	Format
Bank marketing sample data - P4 LGBM Classifier	Model

Select version

 Promoting a version of an asset to a space creates a new asset in the space, with a new asset ID.

Current

Description (optional)

Description of assets

Cancel

Promote



Launch IDE



Find assets

Add asset New asset 

15 assets

All assets 

Asset types

>  Data access 2>  Data 3>  Flows 3>  Visualizations 1

Experiments 2

AutoAI experiment 1

Decision Optimization... 1

>  Source Code 3>  Models 1

Models

Name

Last modified

 Bank marketing sample data - P4 LGBM Classifier Model	18 hours ago Thomas Schaeck (You)
---	--------------------------------------

Successfully promoted Bank marketing sample data - P4 LGBM Classifier to the associated deployment space. Go to the deployment space to prepare the assets for deployment.

Timestamp 1:56:22 PM

Demo for Atruvia

Collaborators

 Thomas Schaeck (You)
Admin

Controls

Restricted eligibility

Cloud object storage

6.5 MB used

IBM Cloud account

Name: Thomas Schäck's Account
ID: f8d190703020c2efd6f009d24c14d71c

SAML federation

IBM w3id

Last modified

9 hours ago

Created on
Apr 01, 2022 by Thomas SchaeckItems per page: 20 

1–1 of 1 items

1 of 1 pages  

Deployments /



Add to space +



Demo Deployment Space

Overview

Assets

Deployments

Jobs

Manage

What assets are you looking for?

Models (1)

Import model +

Name	Type	Software specification	Tags	Last modified	↓
 Bank marketing sample data - P4 LGBM Classifier	wml-hybrid_0.1	hybrid_0.1		Jun 10, 2022 1:56 PM	

Drop files here or browse for files to upload.

Stay on the page until upload completes.
Incomplete uploads are cancelled.

Bank marketing sample data - P4 LGBM Classifier

[New deployment](#) 

Deployments

[Model details](#)

Track this model

The model will be added to your model inventory for activity tracking and model comparison.

 Model tracking available[Track this model](#)  Model inventory Model entry [Track this model](#)Model entry status [Track this model](#) Model informationModel description [Description not added](#)

Tags

Model ID [51f66370-5b65-4a73-9342-a00c2d5b6d75](#)Last modified [Jun 10, 2022, 01:56 PM](#)Created [Jun 10, 2022, 01:56 PM](#)Created by [Thomas Schaeck](#)

Bank marketing sample data - P4 LGBM Classifier

Created [Jun 10, 2022 1:56 PM](#)Type [wml-hybrid_0.1](#)Model ID [51f66370-5b65-4a73-9342-a00c2d5b6d75](#) Software specification [hybrid_0.1](#) Hybrid pipeline software specifications [autoai-kb_rt22.1-py3.9](#)Description [No description provided.](#) Tags [Add tags to make assets easier to find.](#) Source asset details 

Create a deployment

Associated asset

Bank marketing sample data - P4 LGBM Classifier

Deployment type

Online



Run the model on data in real-time, as data is received by a web service.

Batch

Run the model against data as a batch process.

Name

Credit Approval Model Deployment

Serving name 

creditapproval

Description

Deployment description

Tags

Add tags to make assets easier to find.

Cancel

Create



Demo Deployment Space

Overview

Assets

Deployments

Jobs

Manage

What deployments are you looking for?

Deployments (1)

Name	Type	Status	Asset	Tags	Last modified
 Credit Approval Model Deployment	Online	 In progress	Bank marketing sample data - P4 LGBM Classifier		Jun 10, 2022 2:01 PM

Drop files here or browse for files to upload.

Stay on the page until upload completes.
Incomplete uploads are cancelled.

Credit Approval Model Deployment

 Deployed
 Online

API reference

Test

Deployment details

Direct link

Endpoint

<https://us-south.ml.cloud.ibm.com/ml/v4/deployments/creditapproval/predictions?version=2022-06-10>
Bearer <token> 

<https://us-south.ml.cloud.ibm.com/ml/v4/deployments/4babdc73-f78a-430c-a654-03e7fbb0f10d/predictions?version=2022-06-10>


Code snippets

cURL

Java

JavaScript

Python

Scala

```
# NOTE: you must set $API_KEY below using information retrieved from your IBM Cloud account.
```

```
curl --insecure -X POST --header "Content-Type: application/x-www-form-urlencoded" --header "Accept: application/json"
--data-urlencode "grant_type=urn:ibm:params:oauth:grant-type:apikey"
--data-urlencode "apikey=$API_KEY" "https://iam.cloud.ibm.com/identity/token"
```

```
# the above CURL request will return an auth token that you will use as $IAM_TOKEN in the scoring request below
# TODO: manually define and pass values to be scored below
curl -X POST --header "Content-Type: application/json" --header "Accept: application/json" --header "Authorization:
Bearer $IAM_TOKEN" -d '{"input_data": [{"fields": ["$ARRAY_OF_INPUT_FIELDS"], "values": ["$ARRAY_OF_VALUES_TO_BE_SCORED",
$ANOTHER_ARRAY_OF_VALUES_TO_BE_SCORED"]}]}' "https://us-south.ml.cloud.ibm.com/ml/v4/deployments/4babdc73-f78a-430c-a654-03e7fbb0f10d/predictions?version=2022-06-10"
```



Credit Approval Model Deployment

Created

Jun 10, 2022 2:01 PM

Updated

Jun 10, 2022 2:01 PM

Deployment ID

4babdc73-f78a-430c-a654-03e7f... 

Software specification

hybrid_0.1 

Hybrid pipeline software specifications

autoai-kb_rt22.1-py3.9

Copies

1 

Serving name

creditapproval 

Description

No description provided. 

Tags

Add tags to make assets easier to find. 

Associated asset

 Bank marketing sample data - P4 LGBT

51f66370-5b65-4a73-9342-a00c... 

Model ID

51f66370-5b65-4a73-9342-a00c... 

Monitor Model for Fairness, Drift, Accuracy, ...

Insights dashboard

[Refresh](#) [Add to dashboard](#)

Deployments Monitored	Quality Alerts	Fairness Alerts	Drift Alerts	Custom Alerts
3	2	3	0	--

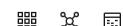
Filter by [Tags](#) Alert type Machine learning provider

Sort by Severity

Which deployment are you looking for?

WOS ExpressPath WML production binding
GermanCreditRiskModel

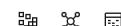
Issues	Quality	1
3	Fairness	2
Drift	--	--
Custom	--	--



Evaluated 2 days ago

WOS ExpressPath WML pre_production bin...
GermanCreditRiskModelChall...

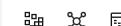
Issues	Quality	--
1	Fairness	1
Drift	--	--
Custom	--	--



Evaluated last month

WOS ExpressPath WML pre_production bin...
GermanCreditRiskModelPrePr...

Issues	Quality	1
1	Fairness	--
Drift	--	--
Custom	--	--



Evaluated last year



Dashboard /

GermanCreditRiskModel

Production

Actions

Last evaluation: Thu, Jun 9, 2022, 4:16 PM CEST

Test data set

--

Number of explanations

1



Tests passed

1

Tests failed

2

OpenPages model

[MOD_0000003](#)

OpenPages model description

RC GCR Model - 24ffaa9f-4597-476a-840c-2c0360c7c000

Fairness



Alerts triggered

Alerts

2

Quality



Alerts triggered

Alerts

1

Drift



No alerts

Alerts

0



Fairness

Alerts triggered

Alerts

! 2

Metric	Score	Violation
Age	75.60%	22.40%
Sex	84.40%	13.60%

Quality

Alerts triggered

Alerts

! 1

Metric	Score	Violation
True positive rate (TPR)	0.62	--
Area under ROC	0.76	0.19
Precision	0.78	--
F1-Measure	0.69	--
Accuracy	0.81	--
Logarithmic loss	0.43	--
False positive rate (FPR)	0.09	--
Area under PR	0.69	--
Recall	0.62	--

Drift

No alerts

Alerts

✓ 0

Metric	Score	Violation
Drop in data consistency	4.16%	none
Drop in accuracy	1.13%	none
Predicted accuracy	78.87%	--
Base accuracy	82.60%	--



Dashboard / GermanCreditRiskModel /

Evaluations

Model ID: f9c1b9cf-104a-4f18-b5b2-57f6305fc90d

Created date: 1/1/1970

[Configure monitors](#)

Fairness

Age	
Sex	
Quality	
Area under ROC	
Area under PR	
Accuracy	
True positive rate (TPR)	
False positive rate (FPR)	
Recall	
Precision	
F1-Measure	
Logarithmic loss	
Drift	
Drop in accuracy	

Fairness for Age

The model's propensity to deliver favorable outcomes to one group over another. [Learn more](#)

Click on a data point to view details.

Time frame	Date range	View
Daily	Custom range	Standard

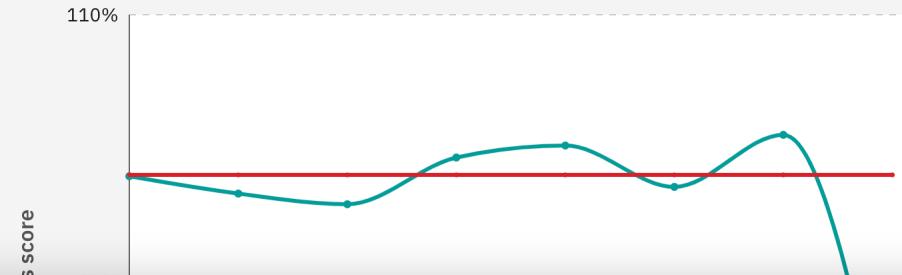
From

10/12/2021 10/19/2021

110%

Score

100%



Mon, Oct 18, 2021, 12:00 AM CEST

Fairness score for Age **101%**

Threshold

98%

Monitored groups

Average	101%
18-25	101%

[Dashboard](#) / [GermanCreditRiskModel](#) /

Evaluations

Model ID: f9c1b9cf-104a-4f18-b5b2-57f6305fc90d

Created date: 1/1/1970

[Configure monitors](#)

Fairness

[Age](#) [Sex](#)

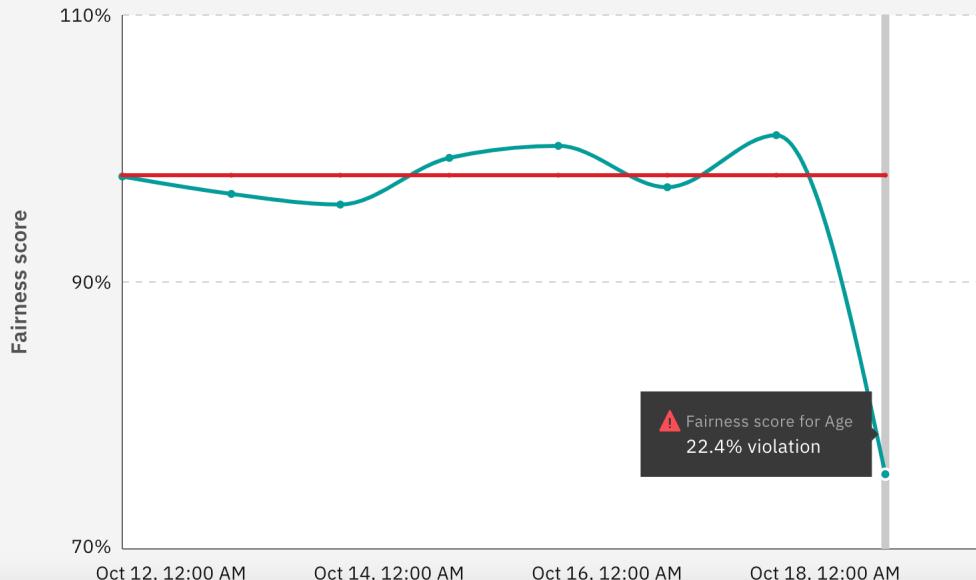
Quality

[Area under ROC](#) [Area under PR](#)[Accuracy](#)[True positive rate \(TPR\)](#)[False positive rate \(FPR\)](#)[Recall](#)[Precision](#)[F1-Measure](#)[Logarithmic loss](#)

Drift

[Drop in accuracy](#)

10/12/2021 10/19/2021



75.6%

Threshold

98%

Monitored groups

Average

75.6%

18-25

75.6%

Schedule

Last evaluation

4:05 PM CEST

Next evaluation

5:05 PM CEST

Violations

(10/12/21-10/19/21)

5

[Evaluate fairness now](#)



Dashboard / GermanCreditRiskModel /

Evaluations

Model ID: f9c1b9cf-104a-4f18-b5b2-57f6305fc90d

Created date: 1/1/1970

[Configure monitors](#)

Fairness

- Age
- Sex

Quality

- [Area under ROC](#)

- Area under PR
- Accuracy
- True positive rate (TPR)
- False positive rate (FPR)
- Recall
- Precision
- F1-Measure
- Logarithmic loss

Drift

- Drop in accuracy

Area under ROC

Area under recall and false positive rate curve. [Learn more](#)

Click on a data point to view details.

Time frame	Date range	View
Daily	Custom range	Standard

From

10/12/2021 10/19/2021

0.75

1

der ROC

0.5

Wed, Oct 13, 2021, 12:00 AM CEST

Area under ROC

0.98

Threshold

0.95

Schedule

Last evaluation 4:05 PM CEST

Next evaluation 5:05 PM CEST

[Evaluate quality now](#)



Dashboard / GermanCreditRiskModel /

Evaluations

Model ID: f9c1b9cf-104a-4f18-b5b2-57f6305fc90d

Created date: 1/1/1970

[Configure monitors](#)

Fairness

Age

Sex

Quality

Area under ROC

Area under PR

Accuracy

True positive rate (TPR)

False positive rate (FPR)

Recall

Precision

F1-Measure

Logarithmic loss

Drift

[Drop in accuracy](#)

Drop in accuracy

The drift monitor estimates the drop in accuracy of the model and the drop in data consistency based on the training data.

Click on a data point to view details.

Time frame	Date range	View
------------	------------	------

Daily

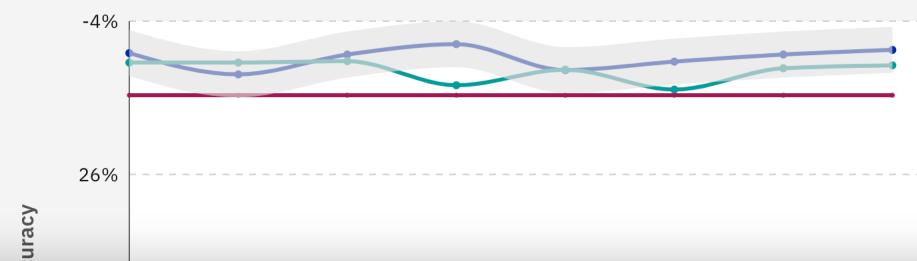
Custom range

Standard

From

To

10/12/2021 10/19/2021



Wed, Oct 13, 2021, 12:00 AM CEST

Drop in accuracy

5.92%

Threshold 10%

Margin of error 1.42% - 10.42%

Drop in data consistency

3.61%

[Dashboard](#) / [GermanCreditRiskModel](#) /

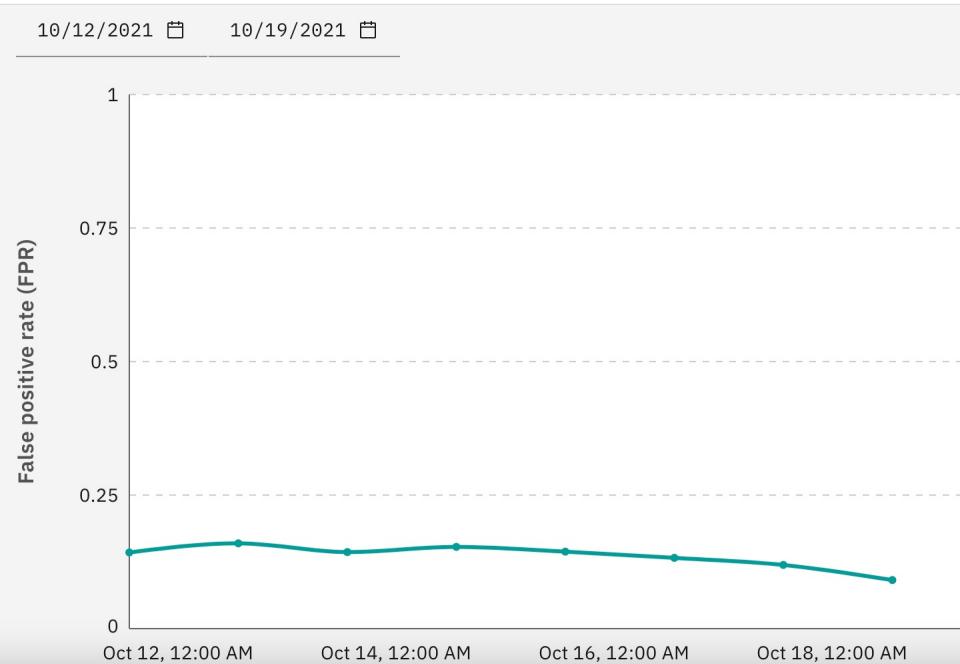
Evaluations

Model ID: f9c1b9cf-104a-4f18-b5b2-57f6305fc90d

Created date: 1/1/1970

[Configure monitors](#)

Age	
Sex	
Quality	
Area under ROC	
Area under PR	
Accuracy	
True positive rate (TPR)	
False positive rate (FPR)	
Recall	
Precision	
F1-Measure	
Logarithmic loss	
Drift	
Drop in accuracy	
Performance	
Number of records	



0.16

Schedule

Last evaluation 5:05 PM CEST

Next evaluation 6:05 PM CEST

[Evaluate quality now](#)[Upload feedback data](#)**Recommendation****If metric is trending upwards**

Metric is deteriorating. Feedback data is becoming significantly different than the training data.

If metric is trending downwards

Metric is improving. Model retraining is



Find a transaction

Deployed model

GermanCreditRiskModel

Recent transactions

 Enter a transaction ID

Transaction ID	Timestamp	Prediction	Confidence	Actions
4b10a990bc313da4f3a8edd1d285ac39-1	Oct 19, 2021, 11:10:20 AM	No Risk	86.37%	Explain
15737345759fb138b7449c21c18acb7c-1	Oct 19, 2021, 11:10:20 AM	Risk	53.89%	Explain
d346dfb591826a197d9defbb896bdd56-1	Oct 19, 2021, 11:10:20 AM	Risk	58.01%	Explain
9fd9214907b587854b642d6c949c3451-1	Oct 19, 2021, 11:10:20 AM	No Risk	53.19%	Explain
66aa36938ee947ad693ca75ee611722c-1	Oct 19, 2021, 11:10:20 AM	No Risk	53.19%	Explain
849e630918ce34106bd55222ba5bc2ae-1	Oct 19, 2021, 11:10:20 AM	Risk	54.31%	Explain

System setup

Prepare Watson OpenScale for use by connecting it to a database, machine learning providers, and integrated services.

Required

- Database
- Machine learning providers
- Users & roles

Optional

- Integrations

Required

Machine learning providers

Description

Watson OpenScale connects to deployed models stored in a machine learning environment.

Add machine learning provider

Watson Machine Learning

WOS ExpressPath WML ...

Watson Machine Learning

WOS ExpressPath WML ...

Description

WML Instance designated as
pre_production

Description

WML Instance designated as
production

Pre-production

Production

System setup

Prepare Watson OpenScale for use by connecting it to a database, machine learning providers, and integrated services.

Required

Database

Machine learning providers

Users & roles

Optional

Integrations

Machine learning providers

Connection

New provider

Description

Connect to the provider where your deployed models are stored and specify if the environment is a pre-production or production environment.

Pre-production environments

Test models by uploading test data sets (csv files) and running evaluations. When the model is ready, approve it for production.

Production environments

Monitor production models by logging model transactions and sending feedback (labeled test data) to Watson OpenScale for continuous evaluation.

Service provider

Choose an option

Watson Machine Learning (V2)

Custom Environment

Amazon SageMaker

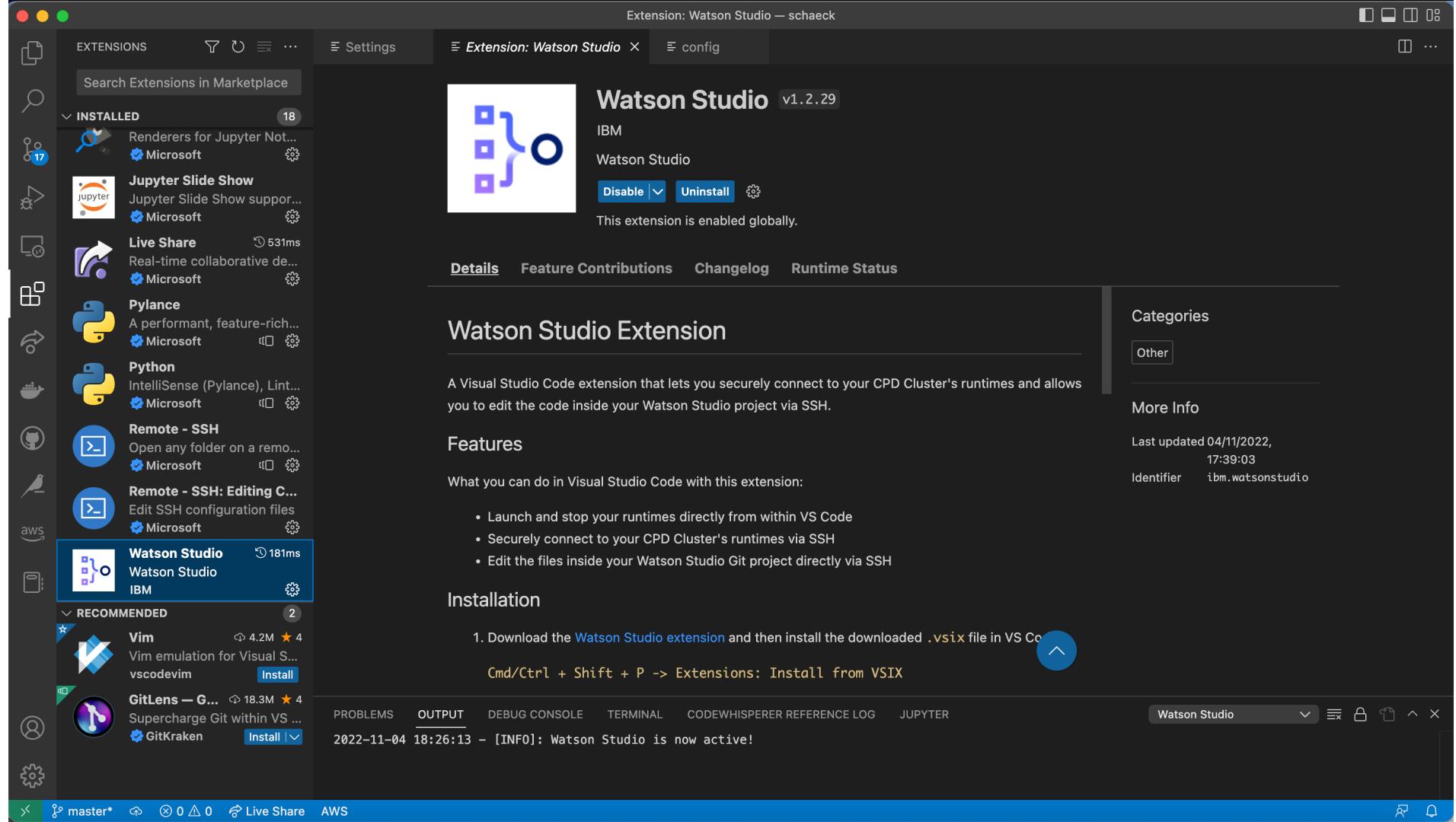
Microsoft Azure ML Studio

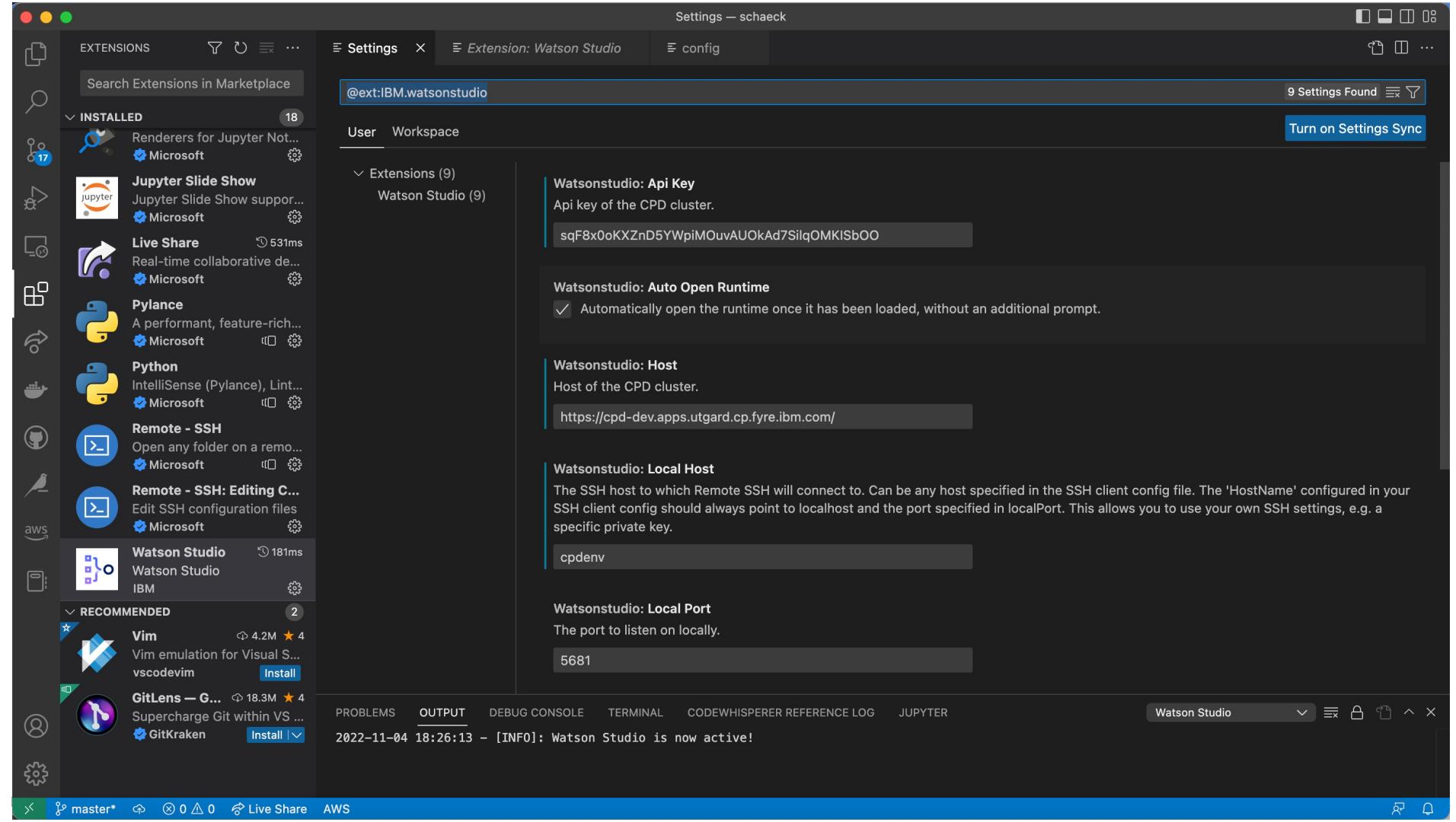
Microsoft Azure ML Service

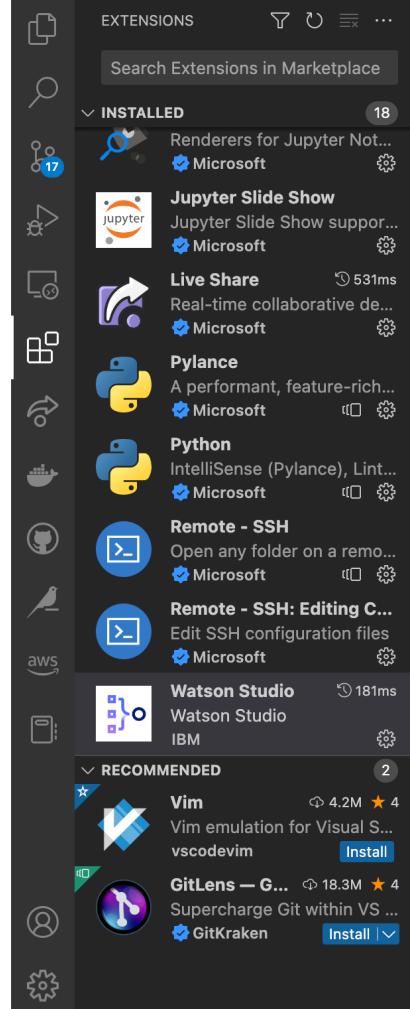
Cancel

Save

Coming in CPD 4.6.0 (November)







Watson Studio: Connect to Runtime

recently used

User Works

Extensions Watson S

Add Browser Breakpoint
Add Cell Tag
Add Function Breakpoint
AWS: About Toolkit
AWS: Add SAM Debug Configuration
AWS: Choose AWS Profile...
AWS: CodeWhisperer Invoke Service
AWS: CodeWhisperer Settings
AWS: Connect to AWS
AWS: Create a new Step Functions state machine

other commands

Host of the CPD cluster.
<https://cpd-dev.apps.utgard.cp.fyre.ibm.com/>

Watsonstudio: Local Host
The SSH host to which Remote SSH will connect to. Can be any host specified in the SSH client config file. The 'HostName' configured in your SSH client config should always point to localhost and the port specified in localPort. This allows you to use your own SSH settings, e.g. a specific private key.
cpdenv

Watsonstudio: Local Port
The port to listen on locally.
5681

9 Settings Found

Turn on Settings Sync

EXTENSIONS

Search Extensions in Marketplace

INSTALLED 12

-  Jupyter keymaps for notebooks by Microsoft
-  Jupyter Notebook Renderers by Microsoft
-  PyLance by Microsoft
-  Python by Microsoft
-  Remote - SSH by Microsoft
-  Remote Development by Microsoft
-  Watson Studio by IBM
-  WSL by Microsoft

RECOMMENDED 5

-  Docker by Microsoft [Install](#)
-  Kubernetes by Microsoft [Install](#)
-  GitLens — Git supercharged by GitLens

Watson Studio v1.2.12

IBM Watson Studio

Disable Uninstall ⚙

This extension is enabled globally.

Details Feature Contributions Changelog Runtime Status

Watson Studio Extension

Categories Other

More Info Identifier ibm.watsonstudio

A Visual Studio Code extension that lets you securely operate your local code with remote Watson Studio resources.

Features

What you can do in Visual Studio Code with this extension:

- Use your local code to access the default .lunvhrlab runtime

PROBLEMS OUTPUT TERMINAL JUPYTER DEBUG CONSOLE

Watson Studio

2022-09-28 10:06:09 - [INFO]: Watson Studio is now active!

2022-09-28 10:07:56 - [INFO]: Authenticating with jschaeck against cpd-dev.apps.brooklyn.cp.fyre.ibm.com

2022-09-28 10:07:58 - [INFO]: Retrieved Git projects



EXTENSIONS ...

Search Extensions in Marketplace

INSTALLED 18

- Renderers for Jupyter Notebooks Microsoft
- Jupyter Slide Show Jupyter Slide Show support... Microsoft
- Live Share Real-time collaborative development Microsoft
- Pylance A performant, feature-rich Python linter Microsoft
- Python IntelliSense (Pylance), Linting Microsoft
- Remote - SSH Open any folder on a remote host Microsoft
- Remote - SSH: Editing Configuration Edit SSH configuration files Microsoft
- Watson Studio Watson Studio IBM

RECOMMENDED 2

- Vim Vim emulation for Visual Studio Code vscodevim
- GitLens — GitHub Git Supercharge Git within VS Code GitKraken

Settings Select the environment to be used.

@ext:IBM.watsonstudio JupyterLab with Runtime 22.1 on Python 3.9
JupyterLab with Runtime 22.2 on Python 3.10 (Running)

User Workspace

Extensions (9) Watson Studio (9)

Watsonstudio: Api Key
Api key of the CPD cluster.
`sqF8x0oKXZnD5YWpiMOuvAUOkAd7SilqOMKISbOO`

Watsonstudio: Auto Open Runtime
 Automatically open the runtime once it has been loaded, without an additional prompt.

Watsonstudio: Host
Host of the CPD cluster.
`https://cpd-dev.apps.utgard.cp.fyre.ibm.com/`

Watsonstudio: Local Host
The SSH host to which Remote SSH will connect to. Can be any host specified in the SSH client config file. The 'HostName' configured in your SSH client config should always point to localhost and the port specified in localPort. This allows you to use your own SSH settings, e.g. a specific private key.
`cpdenv`

Watsonstudio: Local Port
The port to listen on locally.
`5681`

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL CODEWHISPERER REFERENCE LOG JUPYTER

Watson Studio

2022-11-04 18:30:09 - [INFO]: Authenticating with tschaeck against cpd-dev.apps.utgard.cp.fyre.ibm.com
2022-11-04 18:30:10 - [INFO]: Retrieved Git projects
2022-11-04 18:30:23 - [INFO]: Retrieved available environments

IBM Confidential

The screenshot shows the Visual Studio Code interface with the Watson Studio extension installed. The left sidebar displays the Extensions view with various installed and recommended extensions like Jupyter, Python, and Docker. The main panel shows the Watson Studio extension details, including its logo, version (v1.2.12), developer (IBM), and status (enabled globally). The runtime status shows logs from Watson Studio, indicating it is active and authenticating with a service. A progress bar at the bottom indicates the runtime is starting.

Watson Studio v1.2.12

IBM

Watson Studio

Disable Uninstall ⚙️

This extension is enabled globally.

Watson Studio Extension

A Visual Studio Code extension that lets you securely operate your local code with remote Watson Studio resources.

Features

What you can do in Visual Studio Code with this extension:

- Use your local code to access the default .lunvterlab runtime

PROBLEMS OUTPUT TERMINAL JUPYTER DEBUG CONSOLE

Watson Studio

2022-09-28 10:06:09 - [INFO]: Watson Studio is now active!

2022-09-28 10:07:56 - [INFO]: Authenticating with jschaeck against cpd-dev.apps.brooklyn.cp.fyre.ibm.com

2022-09-28 10:07:58 - [INFO]: Retrieved Git projects

2022-09-28 10:08:15 - [INFO]: Retrieved Git projects

2022-09-28 10:08:18 - [INFO]: Retrieved available environments

2022-09-28 10:08:36 - [INFO]: Retrieved Git projects

2022-09-28 10:08:39 - [INFO]: Retrieved available environments

2022-09-28 10:08:54 - [INFO]: Retrieved Git projects

2022-09-28 10:08:57 - [INFO]: Retrieved available environments

2022-09-28 10:08:57 - [INFO]: Starting runtime, this may take a while

Starting runtime, this may take a while

Categories

Other

More Info

Identifier ibm.watsonstudio

Starting runtime

65

The screenshot shows the Visual Studio Code (VS Code) interface with a dark theme. The top bar displays the title "userfs [SSH: vscodehost]". The left sidebar contains icons for Explorer, Open Editors, Search, Issues (with 10), and other development tools. The main editor area shows a Python script named "test2.py". The code imports "itc_utils.flight_service" and uses it to interact with a database, specifically selecting data from the "TEST" table. A context menu is open over the code, with options "Run Python File" and "Debug Python File" visible.

```
1 import itc_utils.flight_service as itcfs
2
3 readClient = itcfs.get_flight_client()
4
5 DB2_data_request = {
6     'connection_name': "DB2",
7     'interaction_properties': {
8         'select_statement': 'SELECT * FROM "SBG64369"."TEST" FETCH FIRST 5000 ROWS ONLY'
9     }
10 }
11
12 flightInfo = itcfs.get_flight_info(readClient, nb_data_request=DB2_data_request)
13
14 data_df_1 = itcfs.read_pandas_and_concat(readClient, flightInfo)
15 print(data_df_1.head(10))
```

Bottom status bar: IBM Cloud, SSH: vscodehost, master*, 0 △ 0, 0, Ln 15, Col 1, Spaces: 4, UTF-8, LF, Python, 3.9.12 ('Python-3.9-Premium': conda), 66

The screenshot shows the Visual Studio Code (VS Code) interface running on a Mac. The title bar indicates the connection is to 'userfs [SSH: vscodehost]'. The left sidebar contains icons for Explorer, Open Editors, UserFS (with 10 items), and other workspace files like test_james.py, test.ipynb, test.py, test2.py, and test4.py. The main editor area displays Python code for interacting with a flight service and reading data from DB2. Below the editor is a terminal window showing the execution of the script and the resulting car data frame.

EXPLORER

OPEN EDITORS

USERFS [SSH: VS CODEHOST]

test2.py

test2.py > ...

```
1 import itc_utils.flight_service as itcfs
2
3 readClient = itcfs.get_flight_client()
4
5 DB2_data_request = {
6     'connection_name': """DB2""",
7     'interaction_properties': {
8         'select_statement': 'SELECT * FROM "SBG64369"."TEST" FETCH FIRST 5000 ROWS ONLY'
9     }
10 }
11
12 flightInfo = itcfs.get_flight_info(readClient, nb_data_request=DB2_data_request)
13
14 data_df_1 = itcfs.read_pandas_and_concat(readClient, flightInfo)
15 print(data_df_1.head(10))
```

PROBLEMS OUTPUT TERMINAL PORTS JUPYTER DEBUG CONSOLE

Python

```
/opt/conda/envs/Python-3.9-Premium/bin/python /userfs/test2.py
● (Python-3.9-Premium) bash-4.4$ conda activate Python-3.9-Premium
● (Python-3.9-Premium) bash-4.4$ /opt/conda/envs/Python-3.9-Premium/bin/python /userfs/test2.py
   MPG CYLINDERS ENGINE HORSEPOWER WEIGHT ACCELERATION YEAR ORIGIN NAME
0 18          8  307.0      130    3504      12.0    70 American chevrolet chevelle malibu
1 15          8  350.0      165    3693      11.5    70 American buick skylark 320
2 18          8  318.0      150    3436      11.0    70 American plymouth satellite
3 16          8  304.0      150    3433      12.0    70 American amc rebel sst
4 17          8  302.0      140    3449      10.5    70 American ford torino
5 15          8  429.0      198    4341      10.0    70 American ford galaxie 500
6 14          8  454.0      220    4354       9.0    70 American chevrolet impala
7 14          8  440.0      215    4312       8.5    70 American plymouth fury iii
8 14          8  455.0      225    4425      10.0    70 American pontiac catalina
9 15          8  390.0      190    3850       8.5    70 American amc ambassador dpl
```

OUTLINE TIMELINE

IBM Corp < SSH: vscodehost master* ↻ ⊞ 0 ▲ 0 ⌂ 0

Ln 15, Col 1 Spaces: 4 UTF-8 LF Python 3.9.12 ('Python-3.9-Premium': conda)

The screenshot shows the Visual Studio Code (VS Code) interface with a dark theme. The top bar displays the title "userfs [SSH: vscodehost]". The left sidebar contains icons for Explorer, Search, Files, Git, Terminal, and Activity, with the "Files" icon having a red notification count of 10. The "EXPLORER" section shows a tree view with "OPEN EDITORS" containing "test2.py", "USERFS [SSH: VS CODEHOST]" containing ".virtual_documents", "assets", "assettypes", ".gitignore", "test_james.py", "test.ipynb", "test.py", "test2.py" (which is currently selected), and "test4.py". The main editor area shows the content of "test2.py":

```
1 import itc_utils.flight_service as itcfs
2
3 readClient = itcfs.get_flight_client()
4
5 DB2_data_request = {
6     'connection_name': "DB2",
7     'interaction_properties': {
8         'select_statement': 'SELECT * FROM "SBG64369"."TEST" FETCH FIRST 5000 ROWS ONLY'
9     }
10 }
11
12 flightInfo = itcfs.get_flight_info(readClient, nb_data_request=DB2_data_request)
13
14 data_df_1 = itcfs.read_pandas_and_concat(readClient, flightInfo)
15 print(data_df_1.head(10))
```

The status bar at the bottom shows "Ln 15, Col 1" and "Spaces: 4". The bottom right corner shows the Python version "3.9.12 ('Python-3.9-Premium': conda)".

This screenshot shows a Python debugging session in VS Code. The interface includes:

- Top Bar:** Shows the title "userfs [SSH: vscodehost]" and standard window controls.
- Left Sidebar:** Includes icons for Run and Debug, Variables, Locals, Globals, Watch, Call Stack, Breakpoints, and Help.
- Call Stack:** Shows the current stack trace: "(module) test2.py:5:1".
- Breakpoints:** Shows breakpoints for "test.ipynb", "test.py", and "test2.py".
- Code Editor:** Displays the file "test2.py" with a breakpoint at line 5. The code uses the `flight_service` API to interact with DB2.
- Terminal:** Shows the command line output of the debugger, indicating it's activated in a "Python-3.9-Premium" environment.
- Output:** Shows the results of the executed code, including the SQL query and its execution.
- Bottom Status Bar:** Provides navigation, search, and status information.

```
test2.py > [e] DB2_data_request
1 import itc_utils.flight_service as itcfs
2
3 readClient = itcfs.get_flight_client()
4
5 DB2_data_request = {
6     'connection_name': """DB2""",
7     'interaction_properties': {
8         'select_statement': 'SELECT * FROM "SBG64369"."TEST" FETCH FIRST 5000 ROWS ONLY'
9     }
10 }
11
12 flightInfo = itcfs.get_flight_info(readClient, nb_data_request=DB2_data_request)
13
14 data_df_1 = itcfs.read_pandas_and_concat(readClient, flightInfo)
15 print(data_df_1.head(10))

(Python-3.9-Premium) bash-4.4$ conda activate Python-3.9-Premium
(Python-3.9-Premium) bash-4.4$ /usr/bin/env /opt/conda/envs/Python-3.9-Premium/bin/python /home/ws
user/.vscode-server/extensions/ms-python.python-2022.14.0/pythonFiles/lib/python/debugpy/adapter/..
///debugpy/launcher 37699 -- /userfs/test2.py
```

A screenshot of the Visual Studio Code interface during a Python debugging session. The title bar shows the workspace name as "userfs [SSH: vscodehost]".

The left sidebar contains several icons: RUN AND DEBUG, VARIABLES, WATCH, CALL STACK, and BREAKPOINTS.

The VARIABLES section shows a variable named "data_df_1" with values: MPG: 12.0, CYLINDERS: 8, and ENGINE: 307.0. It also lists special variables, ACCELERATION: 0, and ENGINE: 0.

The WATCH section is currently empty.

The CALL STACK section shows a single entry: <module> test2.py at line 15:1, with the message "Paused on step".

The main editor area displays the code for "test2.py". The current line being executed is highlighted in yellow: "print(data_df_1.head(10))". Other lines are numbered 1 through 15.

```
1 import itc_utils.flight_service as itcfs
2
3 readClient = itcfs.get_flight_client()
4
5 DB2_data_request = {
6     'connection_name': """DB2""",
7     'interaction_properties': {
8         'select_statement': 'SELECT * FROM "SBG64369"."TEST" FETCH FIRST 5000 ROWS ONLY'
9     }
10 }
11
12 flightInfo = itcfs.get_flight_info(readClient, nb_data_request=DB2_data_request)
13
14 data_df_1 = itcfs.read_pandas_and_concat(readClient, flightInfo)
15 print(data_df_1.head(10))
```

The bottom right corner features a sidebar with tabs for PROBLEMS, OUTPUT, TERMINAL, PORTS, JUPYTER, and DEBUG CONSOLE. The TERMINAL tab is active, displaying a terminal session:

```
(Python-3.9-Premium) bash-4.4$ conda activate Python-3.9-Premium
(Python-3.9-Premium) bash-4.4$ /usr/bin/env /opt/conda/envs/Python-3.9-Premium/bin/python /home/ws
user/.vscode-server/extensions/ms-python.python-2022.14.0/pythonFiles/lib/python/debugpy/adapter/..
/.../debugpy/launcher 37699 -- /userfs/test2.py
```

The bottom status bar shows the file path "SSH: vscodehost", the branch "master*", and other system information like "Ln 15, Col 1" and "70".

SOURCE CONTROL

Message (⌘Enter to commit on 'master')

✓ Commit

Changes

- test2.py (users/test2.py - Modified)
- environment.2coresx4GB.json assets/.METADATA
- hardware_specification.2coresx4GB.json asset...
- job.1fdda2bd-c92a-4636-8f60-f4f450340895...
- job.3223.json assets/.METADATA
- job.867867.json assets/.METADATA
- 2cc0a323-e873-458d-99ea-c26ea740a8bd.lo...
- 7ee65c0d-ef9f-4fcf-b2ab-1edc019f0e31.log a...
- 961eef7d-d616-405c-9d30-4e189ed676d9.lo...
- e31d8952-d18e-49e6-9768-0f0a96526b17.lo...

PROBLEMS OUTPUT TERMINAL PORTS JUPYTER DEBUG CONSOLE

(Python-3.9-Premium) bash-4.4\$ /usr/bin/env /opt/conda/envs/Python-3.9-Premium/bin/python /home/ws user/.vscode-server/extensions/ms-python.python-2022.14.0/pythonFiles/lib/python/debugpy/adapter/.. /.../debugpy/launcher 37699 -- /userfs/test2.py

MPG	CYLINDERS	ENGINE	HORSEPOWER	WEIGHT	ACCELERATION	YEAR	ORIGIN	NAME	
0	18	8	307.0	130	3504	12.0	70	American	chevrolet chevelle malibu
1	15	8	350.0	165	3693	11.5	70	American	buick skylark 320
2	18	8	318.0	150	3436	11.0	70	American	plymouth satellite
3	16	8	304.0	150	3433	12.0	70	American	amc rebel sst
4	17	8	302.0	140	3449	10.5	70	American	ford torino
5	15	8	429.0	198	4341	10.0	70	American	ford galaxie 500
6	14	8	454.0	220	4354	9.0	70	American	chevrolet impala
7	14	8	440.0	215	4312	8.5	70	American	plymouth fury iii
8	14	8	455.0	225	4425	10.0	70	American	pontiac catalina
9	15	8	390.0	190	3850	8.5	70	American	amc ambassador dpl

(Python-3.9-Premium) bash-4.4\$

IBM Cor

SSH: vscodehost master* ↻ ⌂ 0 ▲ 0 ⌂ 0 ⌂ 0

Ln 16, Col 1 Spaces: 4 UTF-8 LF Python 3.9.12 ('Python-3.9-Premium': conda) ⌂ ⌂

The screenshot shows the Visual Studio Code (VS Code) interface with the following components:

- Source Control:** On the left, it displays a list of staged and changes. A commit message "change123" is entered in the commit field, and a "Commit" button is visible.
- Editor:** The main area shows a Python file named "test2.py". The code imports "itc_utils.flight_service" and uses it to get a flight client, then performs a DB2 query and processes flight info.
- Terminal:** At the bottom, the terminal window shows the output of running the script. It includes the command used and the first few rows of the data frame.
- Python Extension:** A sidebar on the right shows the "Python" and "Python De..." sections of the extension's settings.

```
import itc_utils.flight_service as itcfs
readClient = itcfs.get_flight_client()
DB2_data_request = {
    'connection_name': """DB2""",
    'interaction_properties': {
        'select_statement': 'SELECT * FROM "SBG64369"."TEST" FETCH FIRST 5000 ROWS ONLY'
    }
}
flightInfo = itcfs.get_flight_info(readClient, nb_data_request=DB2_data_request)
data_df_1 = itcfs.read_pandas_and_concat(readClient, flightInfo)
print(data_df_1.head(10))
```

```
(Python-3.9-Premium) bash-4.4$ /usr/bin/env /opt/conda/envs/Python-3.9-Premium/bin/python /home/ws/user/.vscode-server/extensions/ms-python.python-2022.14.0/pythonFiles/lib/python/debugpy/adapter/../debugpy/launcher 37699 -- /userfs/test2.py
   MPG CYLINDERS ENGINE HORSEPOWER WEIGHT ACCELERATION YEAR ORIGIN NAME
0 18          8  307.0     130    3504      12.0    70 American chevrolet chevelle malibu
1 15          8  350.0     165    3693      11.5    70 American buick skylark 320
2 18          8  318.0     150    3436      11.0    70 American plymouth satellite
3 16          8  304.0     150    3433      12.0    70 American amc rebel sst
4 17          8  302.0     140    3449      10.5    70 American ford torino
5 15          8  429.0     198    4341      10.0    70 American ford galaxie 500
6 14          8  454.0     220    4354       9.0    70 American chevrolet impala
7 14          8  440.0     215    4312       8.5    70 American plymouth fury iii
8 14          8  455.0     225    4425      10.0    70 American pontiac catalina
9 15          8  390.0     190    3850       8.5    70 American amc ambassador dpl

```

