

Watson Studio

Novák Márton

2019. május 14.

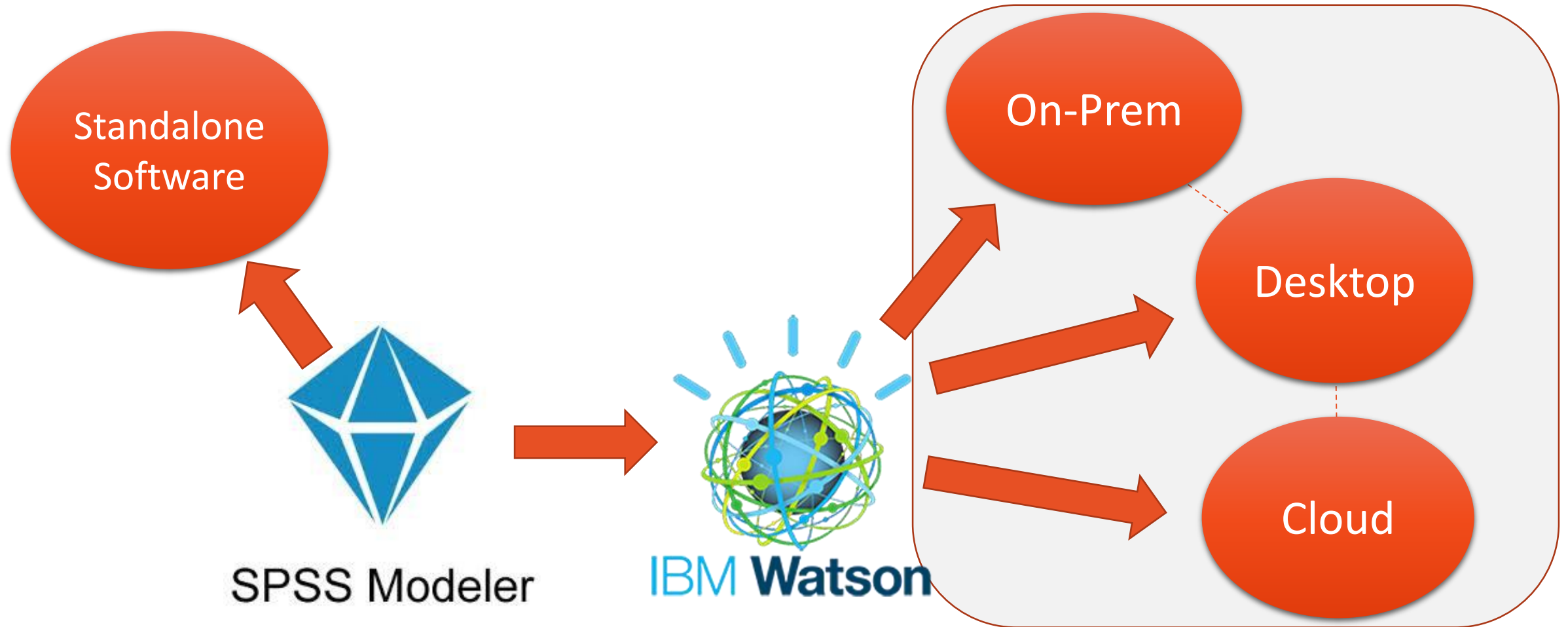
<https://www.ibm.com/cloud/watson-studio>

<https://datascience.ibm.com>

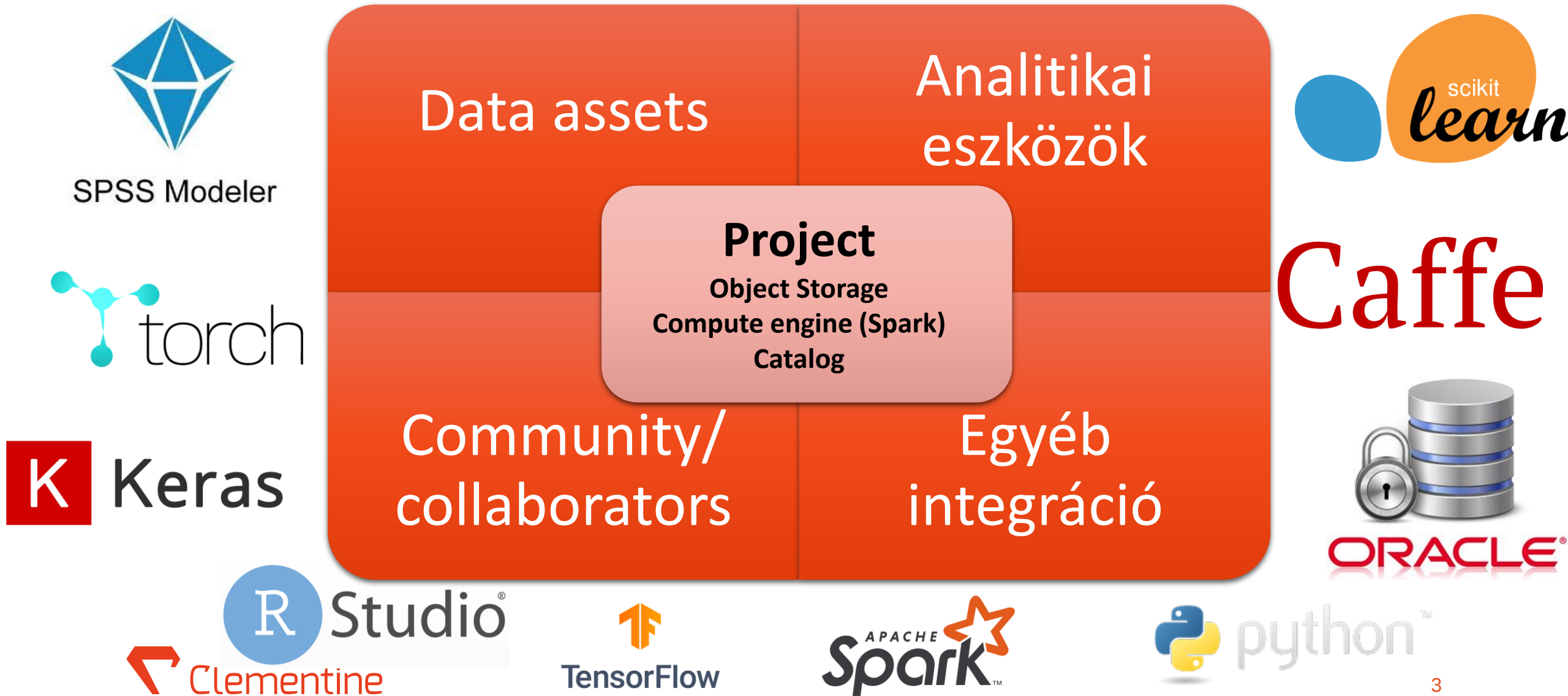
<https://dataplatform.ibm.com>



SPSS MODELER



WATSON STUDIO



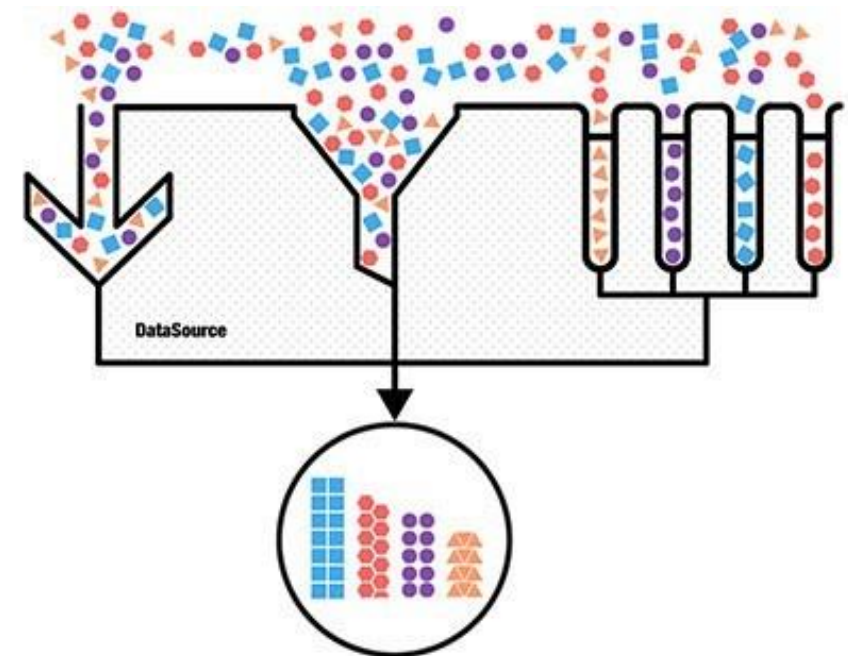
DATA ASSETS

- **Adat file feltöltése object storage-be**
- **Kapcsolat adatforrással**
 - Cloud data services (Db2 Warehouse on Cloud, Amazon S3, Cloudant, PostgreSQL)
 - On-premises databases (Db2, Netezza, Oracle, and SQLServer)
 - Streaming data services (IBM MessageHub and IBM Streaming Analytics)
- **Community data sets**

Search results (304)

DATA SET	
UCI: Automobile Data Set	
AUTHOR IBM	DATE Mar 15, 2019
TOPIC Transportation	
❤️ 2 📌 +	

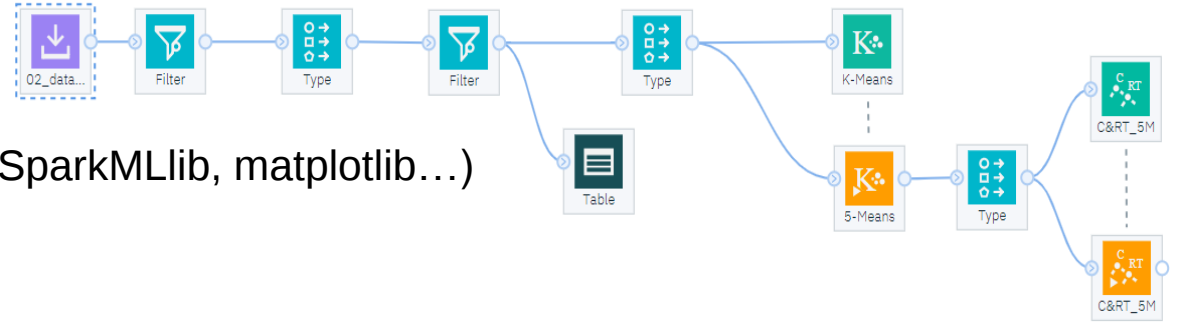
DATA SET	
UCI: SMS Spam Collection - NeuNetS...	
AUTHOR IBM	DATE Dec 14, 2018
TOPIC Society	
❤️ 6 📌 +	



ANALITIKAI ESZKÖZÖK

- **Jupyter notebook:**

- Python/ R/ Scala
- Open source könyvtárak/csomagok (pyspark, SparkMLlib, matplotlib...)
- Hozzáadhatunk hiányzó csomagokat
- Sparkos működés



- **Modeler Flow Editor**

- IBM SPSS Modeler
- Neural Network Modeler
- Spark Mllib

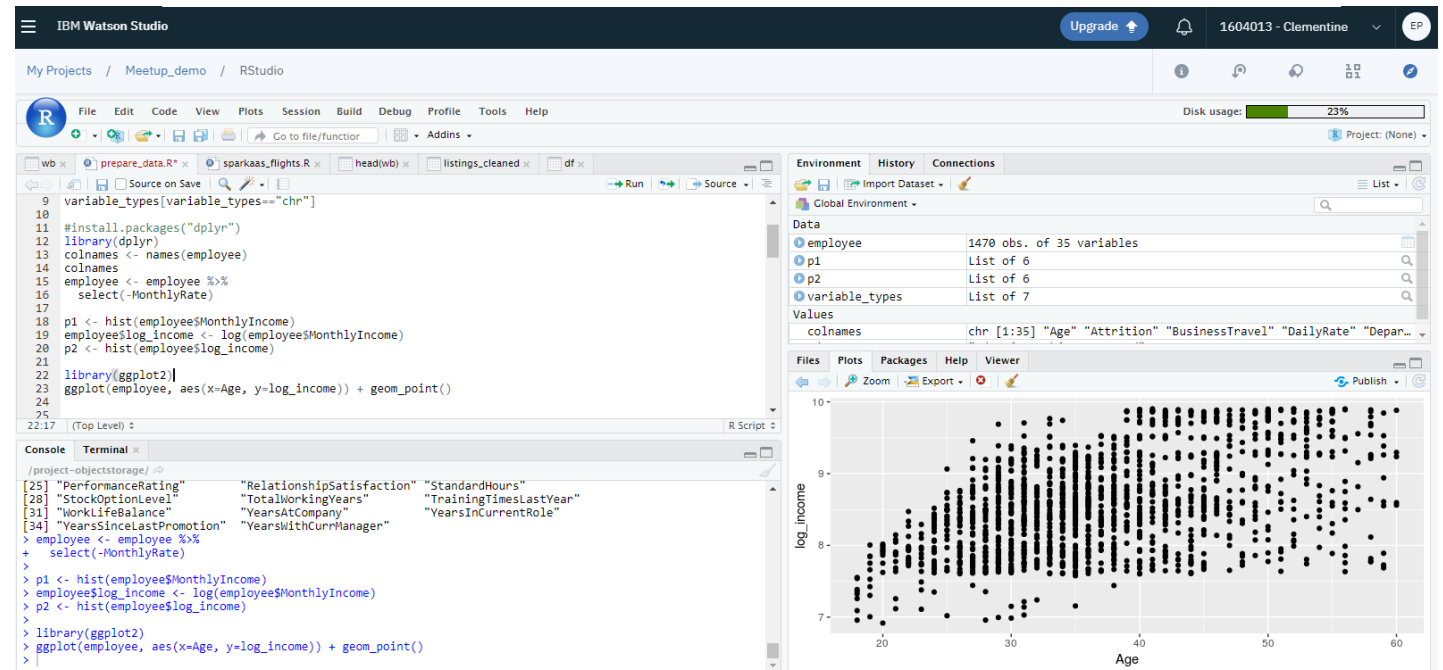
- **Rstudio**

- Shiny Web Applications

- **Experiment Builder**

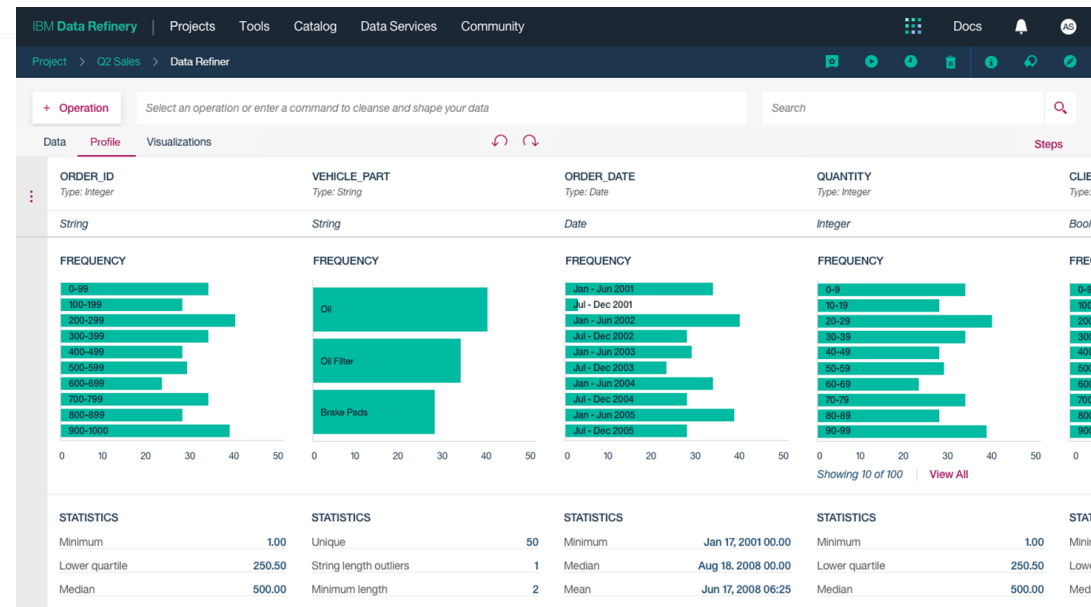
- Machine Learning model tanítás

- **Analytics Dashboard**



EGYÉB INTEGRÁCIÓ

- **Visual Recognition**
 - IBM cloud service
- **Döntés optimalizáció**
 - CPLEX engine
- **Streaming Analytics (Streams Flow)**
 - Real time adatforrások
- **Natural Language Understanding**
- **Natural Language Classifier**
- **Data Refinery**
 - Adatok tisztítása



PROJEKT LÉTREHOZÁSA

The screenshot shows the IBM Watson Studio interface. At the top, a dark blue header contains the 'IBM Watson Studio' logo on the left, an 'Upgrade' button with an upward arrow, a notification bell, the user's name '1604013 - Clementine' with a dropdown arrow, and a profile icon labeled 'MN'. Below the header, a light blue bar contains a 'Get started' link with a dropdown arrow. The main content area has a dark blue background with white text and icons. On the left, it says 'Welcome Márton!' followed by 'Watson Studio • Watson Knowledge Catalog'. Below this, a section titled 'Start by creating a project' explains that a project is for organizing resources. At the bottom, two white boxes offer 'Create a project' and 'Search a catalog' options. On the right, a large white circular graphic contains a line-art illustration of a computer monitor displaying binary code, a magnifying glass, and a document.

IBM Watson Studio

Upgrade ↑

1604013 - Clementine

MN

Get started ▾

Welcome Márton!

Watson Studio • Watson Knowledge Catalog

Start by creating a project

A project is how you organize your resources to work with data and collaborate with team members

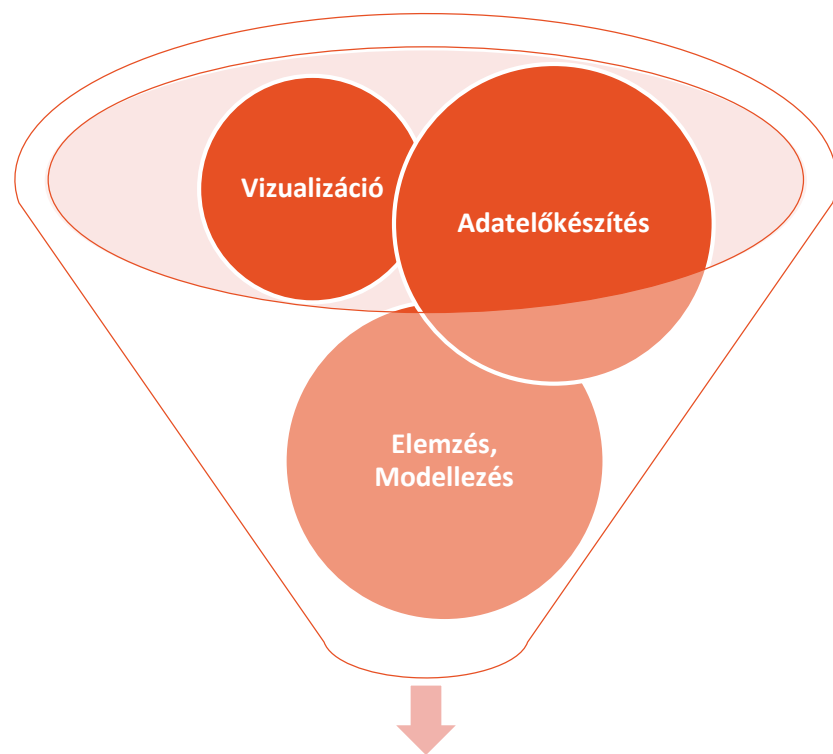
Create a project

Create a project, then add the tools and assets you need.

Search a catalog

Find the assets you need in a catalog.

ESETTANULMÁNY: CREDIT SCORING

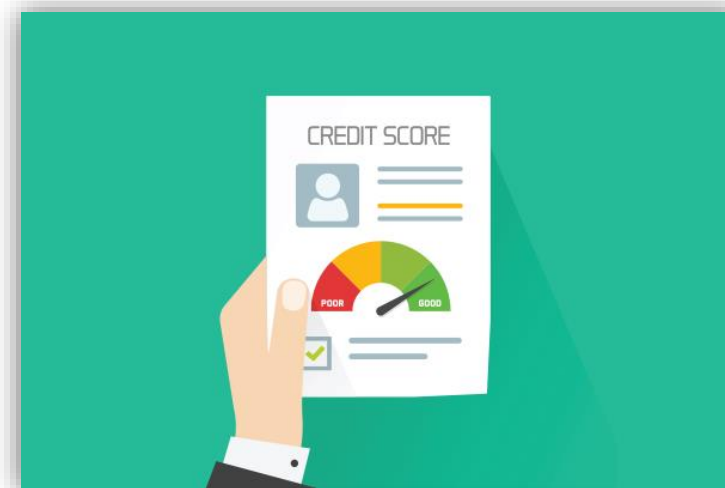


Watson Studio

- Data Refinery
- Notebooks (R/Python/Scala +Spark, TensorFlow)
- R Studio
- Modeler Flow



- **50,000 ügyfél adata**¹
- **Közel 30 magyarázó változó**
 - demográfiai
 - pénzügyi
 - egyéb
- **Bináris célváltozó ~ visszafizeti-e (80% – 20%)**
- **3. helyezett: AUC - 0.655**



1. Linhart, C., Harari, G., Abramovich, S., & Buchris, A. PAKDD data mining competition: New ways of using known methods. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (pp. 99-105). Springer, Berlin, Heidelberg.

ADAT FELFEDEZÉSE, ELŐKÉSZÍTÉSE

Data refinery

The screenshot displays the IBM Watson Studio web interface. At the top, a dark blue header bar contains the 'IBM Watson Studio' logo on the left, and 'Upgrade', a notification bell, the user profile '1604013 - Clementine', and a circular icon with 'MN' on the right. Below the header, a light blue navigation bar shows 'My Projects / otp_proj' on the left and a series of icons on the right, including a dropdown menu, a person icon, 'Launch IDE', and a prominent blue 'Add to project' button. A secondary navigation bar below this contains tabs for 'Overview' (selected), 'Assets', 'Environments', 'Bookmarks', 'Access Control', and 'Settings'. The main content area is divided into two columns. The left column, titled 'otp_proj', shows 'Last Updated: Feb 28 2019' and a 'Readme' link. The right column features three large statistics: '5 Assets', '0 Bookmarks', and '1 Collaborators'. Below these, the 'Recent activity' section is mostly empty, with a dashed box and a message: 'Alerts related to this project will show here when the project is active.' The left sidebar contains several sections: 'Date created' (Feb 28 2019), 'Description' (No description available), 'Storage' (Cloud Object Storage, 12.72 MB used), 'Collaborators' (listing 'Márton Novák Admin' with a 'View all (1)' link), and 'Bookmarks' (showing 'You currently have 0 bookmarks' with a 'View all (0)' link). At the bottom of the interface, a light blue bar contains the 'Readme' title on the left and a 'Back to top' link on the right.

IBM Watson Studio

Upgrade

1604013 - Clementine

MN

My Projects / otp_proj

Launch IDE

Add to project

Overview Assets Environments Bookmarks Access Control Settings

otp_proj

Last Updated: Feb 28 2019

Readme

5 Assets 0 Bookmarks 1 Collaborators

Recent activity

Alerts related to this project will show here when the project is active.

Date created
Feb 28 2019

Description
No description available

Storage
Cloud Object Storage
12.72 MB used

Collaborators
View all (1)
Márton Novák
Admin

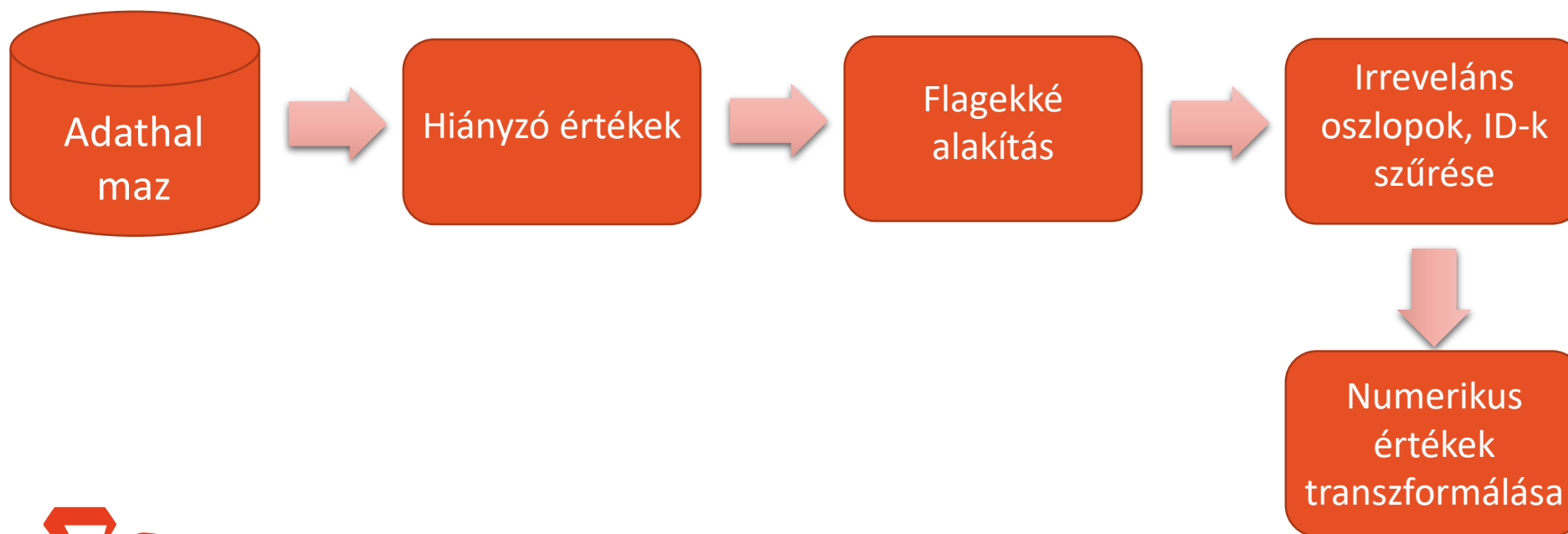
Bookmarks
View all (0)
You currently have 0 bookmarks

Readme

Back to top

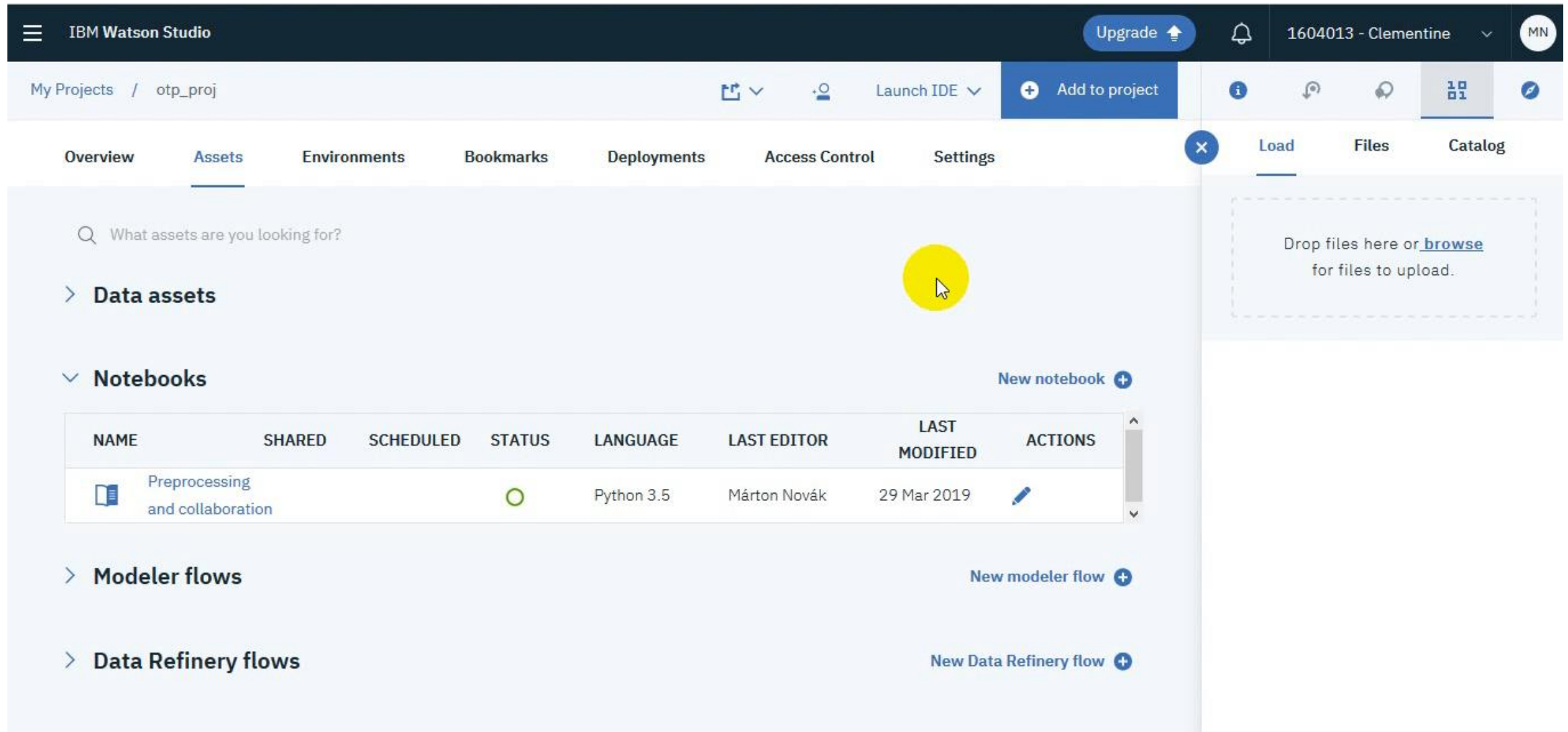
ADATTISZTÍTÁS PYTHONBAN

- Adatok betöltése Pythonba
- Hiányzó értékek szűrése
- Bináris flag-ek konvertálása
- ID-k, irreveláns oszlopok törlése
- Numerikus változók értékeinek binelése, transzformációja



ADATELŐKÉSZÍTÉS ÉS -VIZUALIZÁCIÓ

Python Notebook

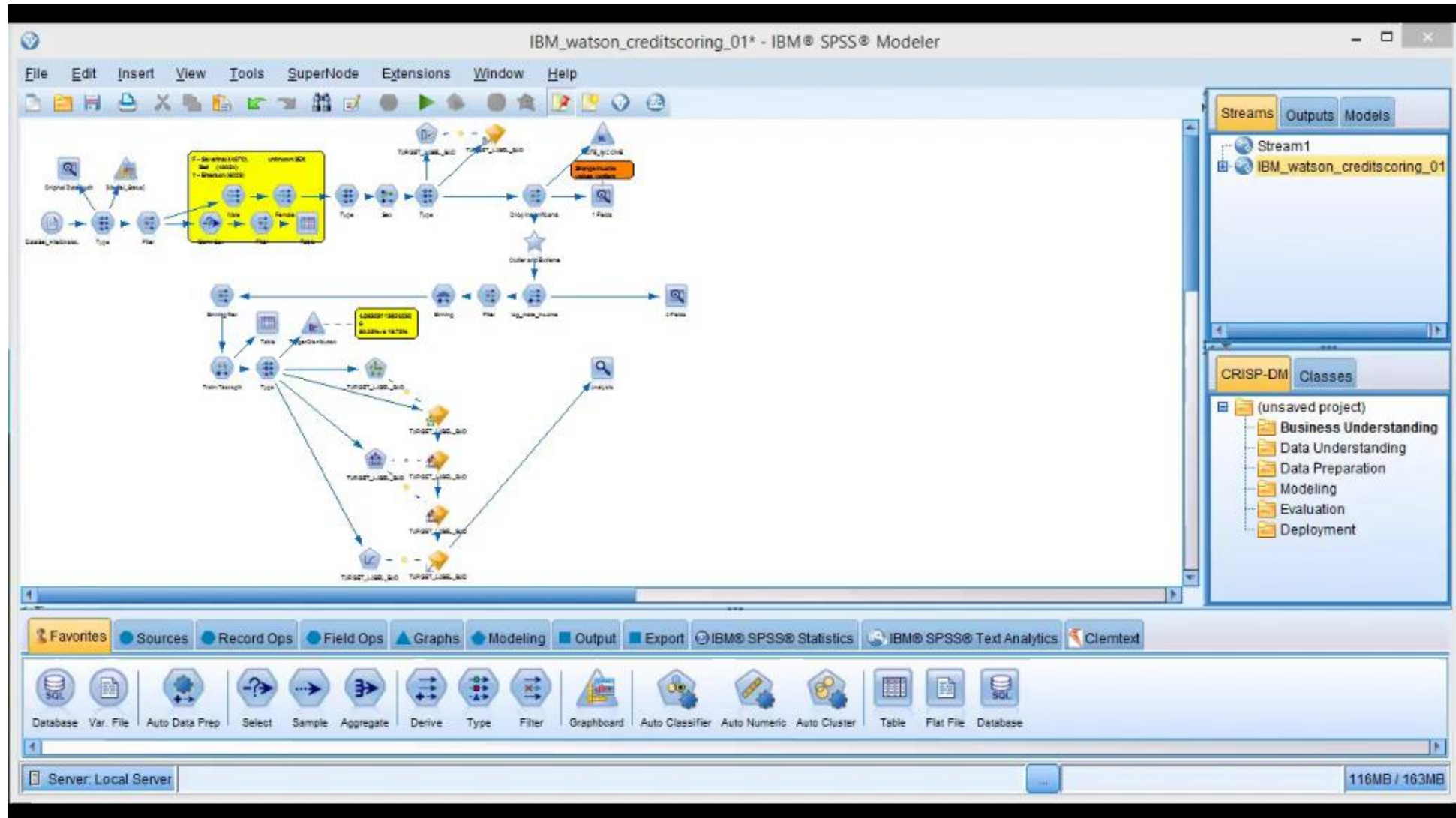


The screenshot displays the IBM Watson Studio web interface. At the top, the header includes the IBM Watson Studio logo, an 'Upgrade' button, a notification bell, the user name '1604013 - Clementine', and a profile icon 'MN'. Below the header, the breadcrumb navigation shows 'My Projects / otp_proj'. The main navigation bar contains tabs for 'Overview', 'Assets' (which is selected), 'Environments', 'Bookmarks', 'Deployments', 'Access Control', and 'Settings'. On the right side of the 'Assets' tab, there are buttons for 'Add to project', 'Load', 'Files', and 'Catalog'. The main content area of the 'Assets' tab features a search bar with the placeholder text 'What assets are you looking for?'. Below the search bar, there are expandable sections for 'Data assets', 'Notebooks', 'Modeler flows', and 'Data Refinery flows'. The 'Notebooks' section is currently expanded, showing a table of notebooks. A yellow circle with a mouse cursor is positioned over the 'Notebooks' section header. To the right of the main content area, a dashed box indicates a file upload area with the text 'Drop files here or [browse](#) for files to upload.'

NAME	SHARED	SCHEDULED	STATUS	LANGUAGE	LAST EDITOR	LAST MODIFIED	ACTIONS
 Preprocessing and collaboration				Python 3.5	Márton Novák	29 Mar 2019	

ELEMZÉS, MODELL ÉPÍTÉS

Modeler flow



NEURAL NETWORKS

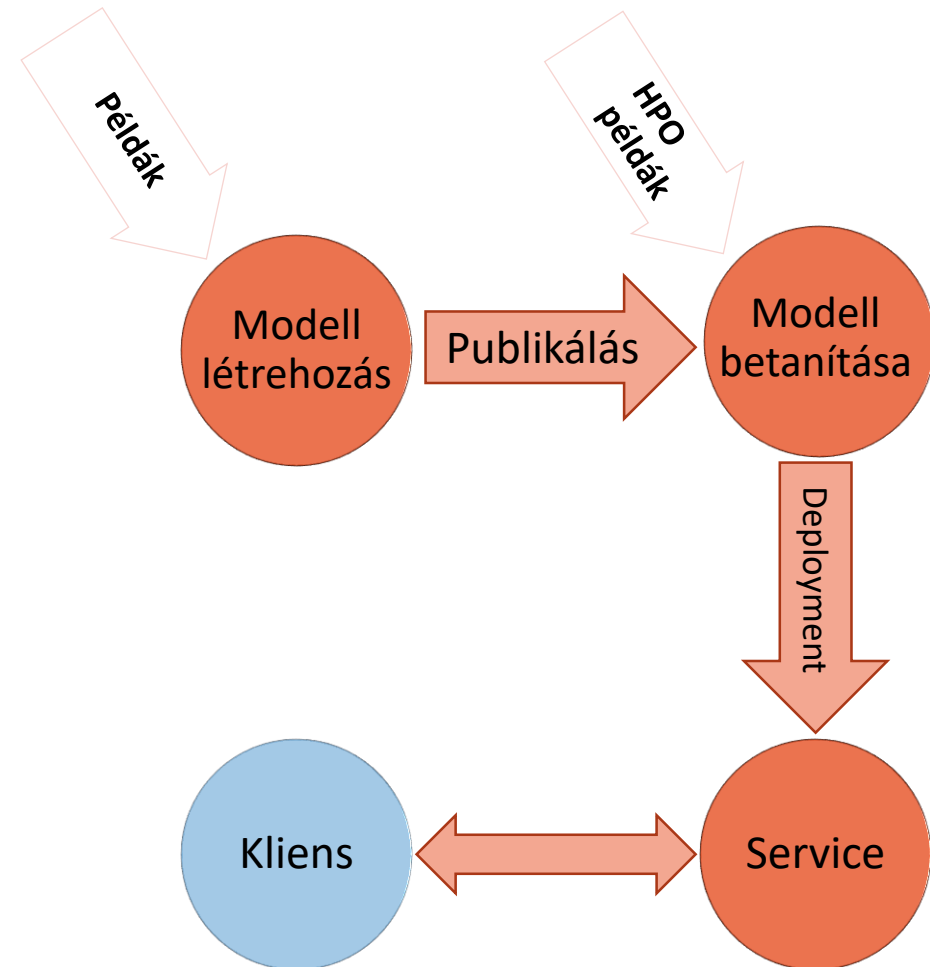
- **Modeler-hez hasonló felület, teljesen más elemek**

- **Elkészült hálózatok exportálása**

- Flow
- TensorFlow
- Keras
- Caffe
- PyTorch

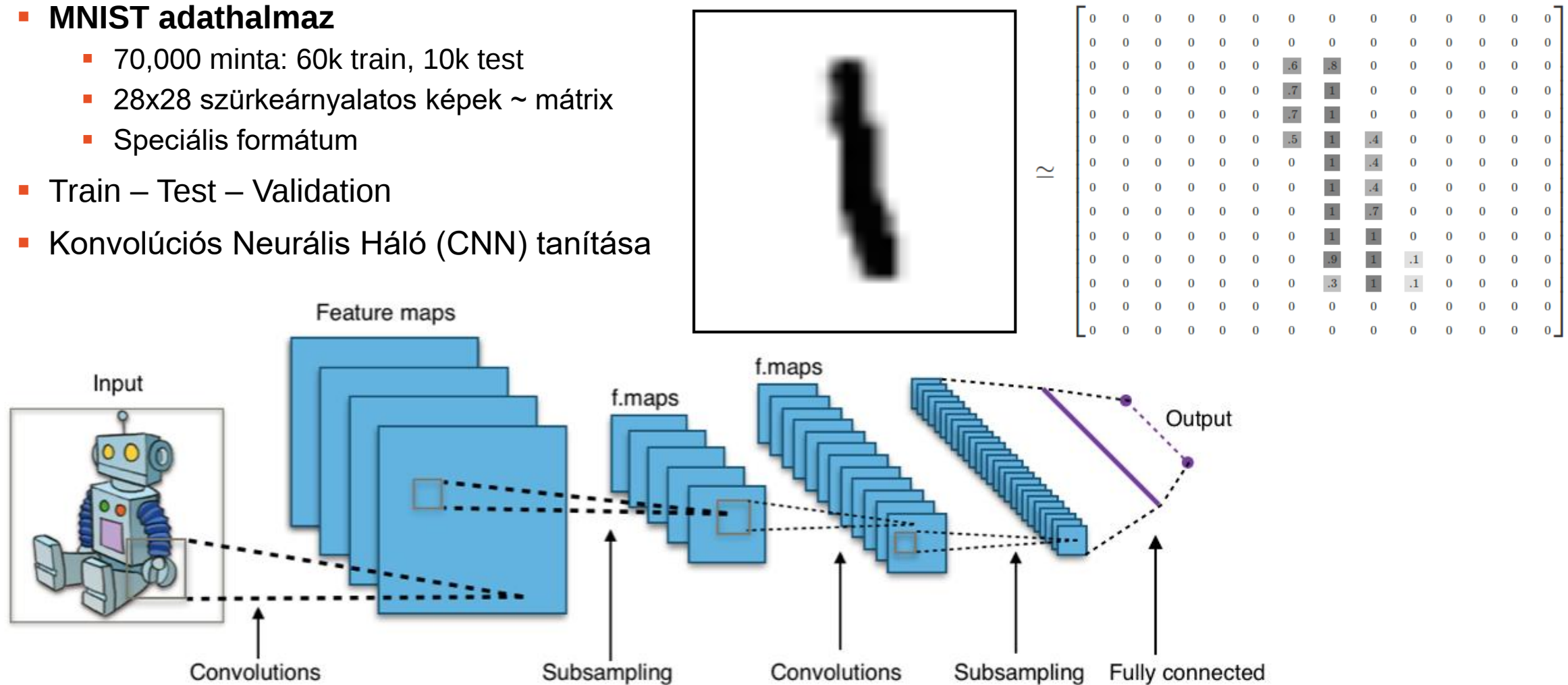
- **Publikálás és Deployment**

- WebService
- Batch előrejelzés (esemény/ütemezés)
- Virtuális Container + file-ok
 - App-ba ágyazás
jelenleg csak CoreML



ESETTANULMÁNY: KÉZZEL ÍRT SZÁMJEGYEK FELISMERÉSE

- **MNIST adathalmaz**
 - 70,000 minta: 60k train, 10k test
 - 28x28 szürkeárnyaltos képek ~ mátrix
 - Speciális formátum
- Train – Test – Validation
- Konvolúciós Neurális Háló (CNN) tanítása



NEURAL NETWORK

The screenshot shows the IBM Watson Studio interface. At the top, a dark blue header bar contains the IBM Watson Studio logo on the left, an 'Upgrade' button with an upward arrow in the center, a notification bell icon, the user's name '1604013 - Clementine' with a dropdown arrow, and a circular profile icon labeled 'MN' on the right. Below the header, a light blue bar features a 'Get started' link with a dropdown arrow. The main content area has a large blue background with white text and a line-art illustration of a computer monitor, a magnifying glass over a binary code circle, and a document icon. The text 'Welcome Márton!' is prominently displayed, followed by 'Watson Studio • Watson Knowledge Catalog'. Below this, a section titled 'Start by creating a project' explains that a project is used to organize resources for working with data and collaborating with team members. At the bottom, two white cards are visible: 'Create a project' (with the subtext 'Create a project, then add the tools and assets you need.') and 'Search a catalog' (with the subtext 'Find the assets you need in a catalog.').

IBM Watson Studio

Upgrade

1604013 - Clementine

MN

Get started

Welcome Márton!

Watson Studio • Watson Knowledge Catalog

Start by creating a project

A project is how you organize your resources to work with data and collaborate with team members

Create a project

Create a project, then add the tools and assets you need.

Search a catalog

Find the assets you need in a catalog.

ÖSSZEGZÉS

- Közös felület egy modellezésen dolgozó minden munkatárs számára
- Könnyű elérést biztosít az erőforrásokhoz
- Modeler integráció és kompatibilitás
- Felhő alapú erőforrás tárolás, elosztott megoldások
- Könnyű, gyors, akár kódmentes fejlesztés
- Erős támogatás
- IBM Watson Studio desktop ~ lokális futtatás
- Legtöbb funkció ingyenesen kipróbálható!



Köszönöm a figyelmet!

Kérdések?