Data Science as a Service

The Next Frontier in Analytics Evolution

Filipa Peleja

Lead Data Science Engineer @ LEGO Group

Lead Data Science Instructor @ CodeOp

*CodeOp – Bridging technical talent development with DE&I through our **international tech school for women+***

Data Science as a Service

Enable data democratization

- Enable everybody in an organization to work with data comfortably and make data-informed decisions

Enable data science as a service

- Enable technology that can be used by everybody to investigate data, build reports, data science solutions

Business Analytics, Data Analytics, and Data Science
are interconnected with **Data Democratization**

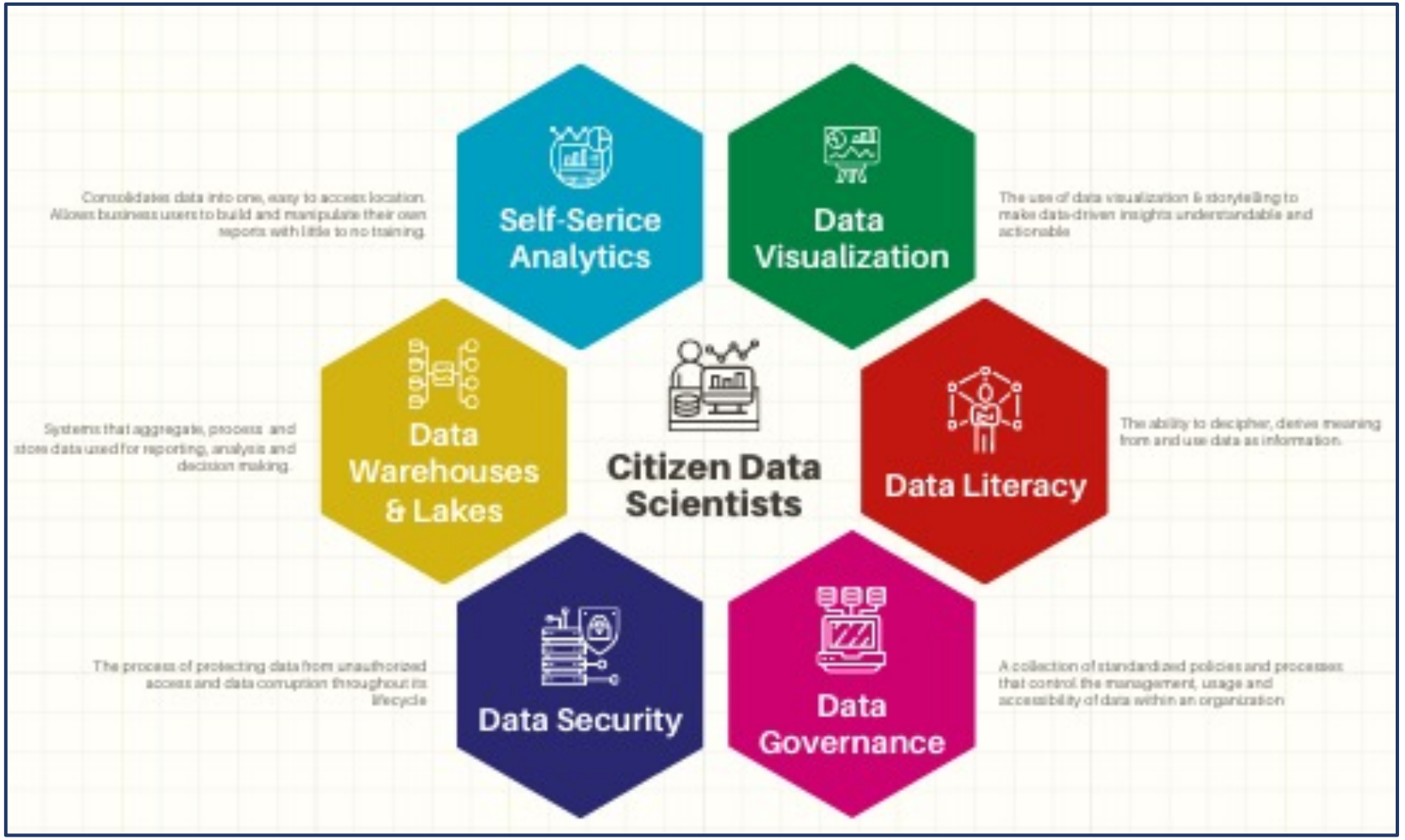
Data Science as a Service

Why is it important?

Data Science as a Service enables businesses to plug and play to start seeing a return on investment fast

Data Science as a Service: who will use it?

Expert & Citizen Data Scientists



https://www.aidataanalytics.network/data-democratization/articles/making-data-accessible-to-everyone-with-data-democratization

Expert vs Citizen Data Scientists

Expert Data Scientists	Citizen Data Scientists
Hold the title and focus on data science solutions	Typically business analysts or data-savvy business users
Studies in data science, statistics, mathematics or another relevant area	Variety of educational backgounds usually with no formal data science studies
Often in **centralised data teams**, and collaborate with specific business units	Part of business unit with **domain-specific knowledge**
Advanced technical skills, write code and create algorithms and analytical models	**Rely on augmented analytics and automated machine learning tools**
Handle **complex analytics** applications and handle **large volumes of data**	**Handle less advanced analytics** and sometimes work with expert data scientists

https://cdn.ttgtmedia.com/rms/onlineimages/expert_data_scientists_vs_citizen_data_scientists-f_mobile.png

Data Science as a Service

Risks - Citizen Data Scientists

> ***" First, auto-ML does not solve for gaps in expertise, training, and experience,* thus increasing the probability of failure.**
>
> ***" AI can go technically or functionally sideways, and non-data scientists with auto-ML may run straight into those pitfalls.***
>
> Reid Blackman and Tamara Sipes (2022) Harvard Business Review

- Handling bias or unbalanced problems
  - *e.g., subsampling, oversampling or tailoring data to a given domain knowledge*
- Knowing if data used is trustworthy, available overtime or choosing a comparable data to evaluate model performance
- Knowing the legal usage of the data for a specific problem as well as ethical, reputational and regulatory risks

https://hbr.org/2022/12/the-risks-of-empowering-citizen-data-scientists

Data Science as a Service

How to mitigate risks?

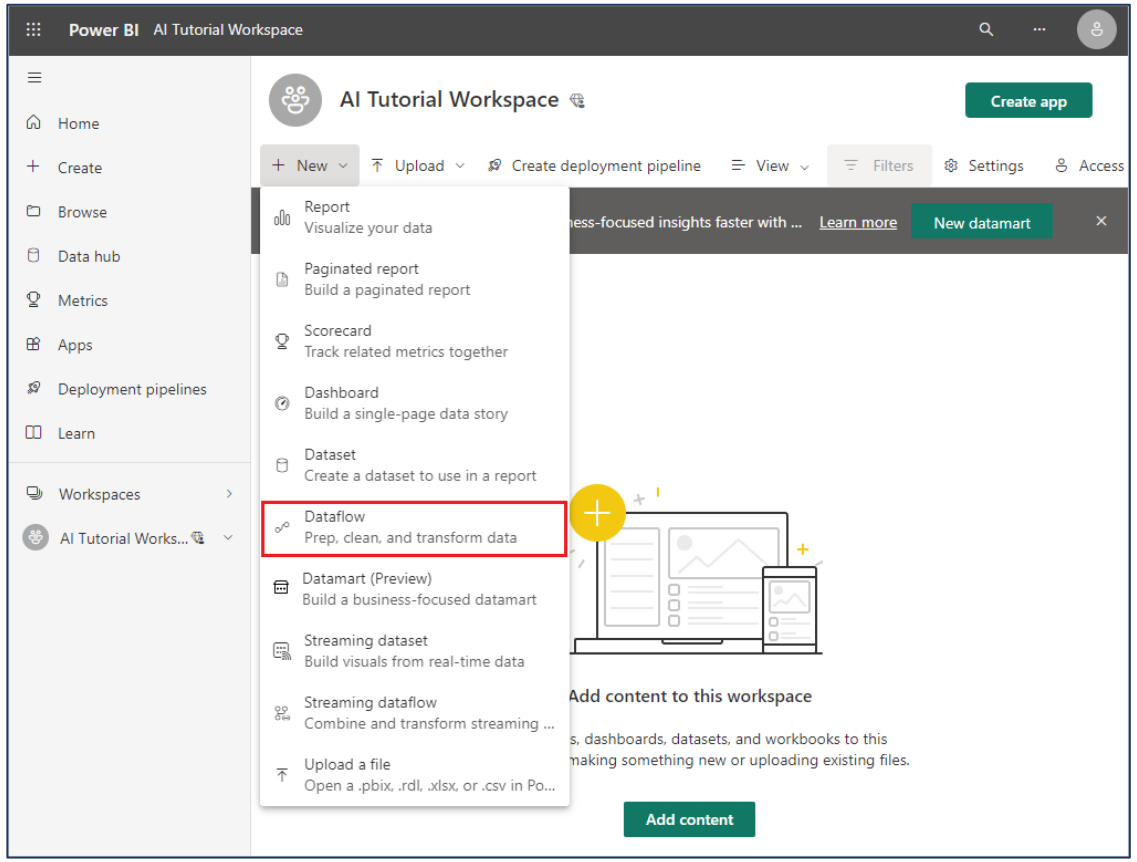
- Create, maintain, and scale **appropriate oversight and guidance**
- Promote continuous learning for Citizen Data Scientists by having access to **published best practices and guidelines**
- **Provide examples or case studies** they can use as **templates** for their own projects
- Promote **collaboration** between Expert & Citizen Data Scientists

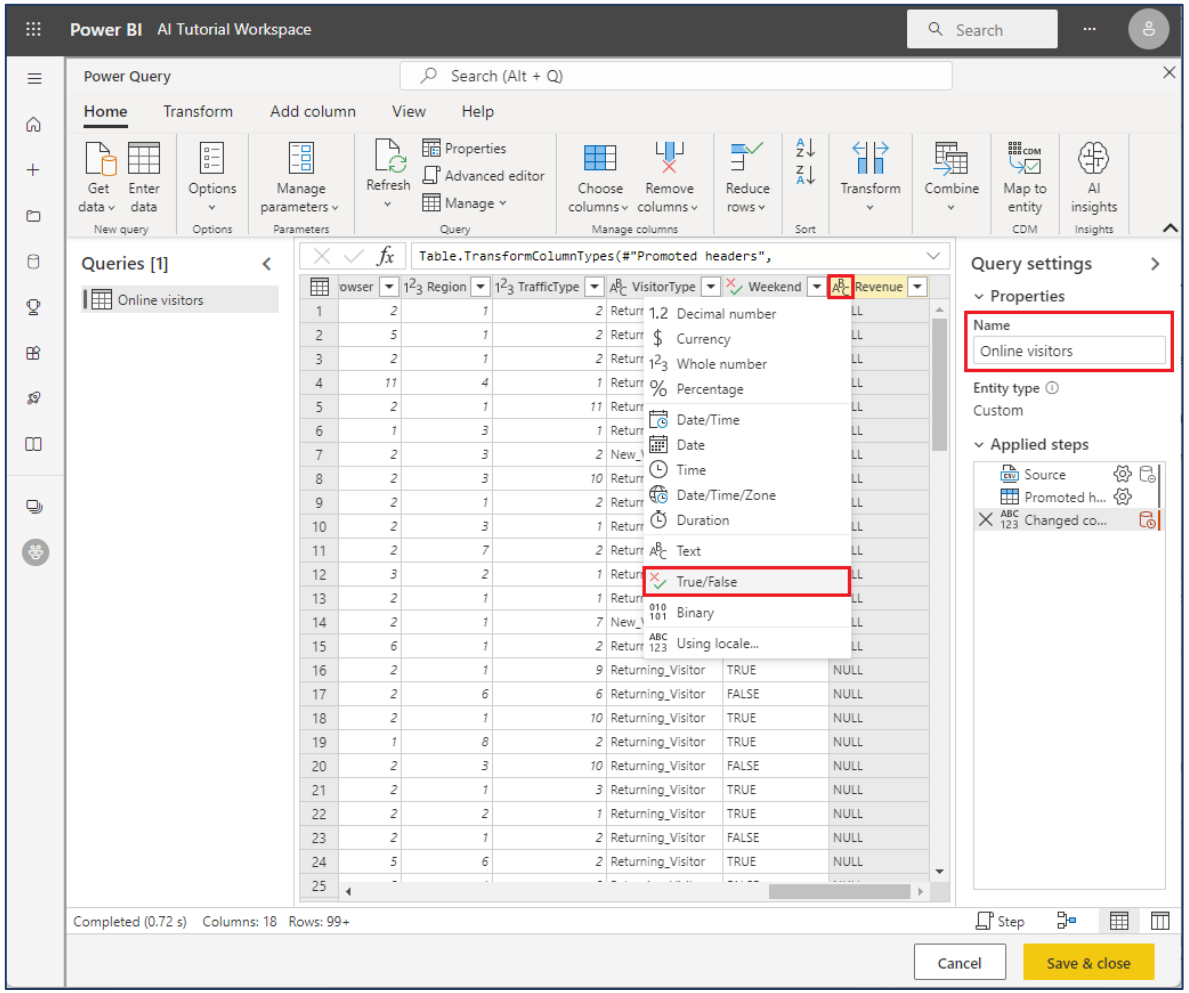
- Verify all projects before moving to production
  - *Need to be assessed and approved for deployment*
  - *Should be comprised of technologists and people from risk, compliance, legal, and ethics*

https://hbr.org/2022/12/the-risks-of-empowering-citizen-data-scientists

Technology for Data Science as a Service

Example: Power BI

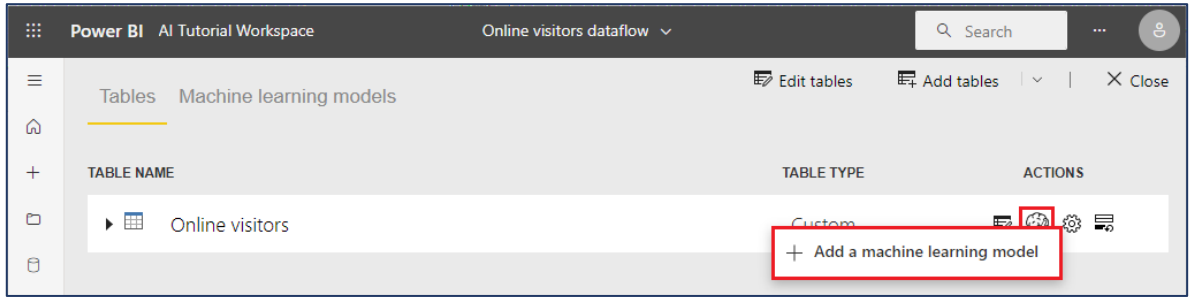


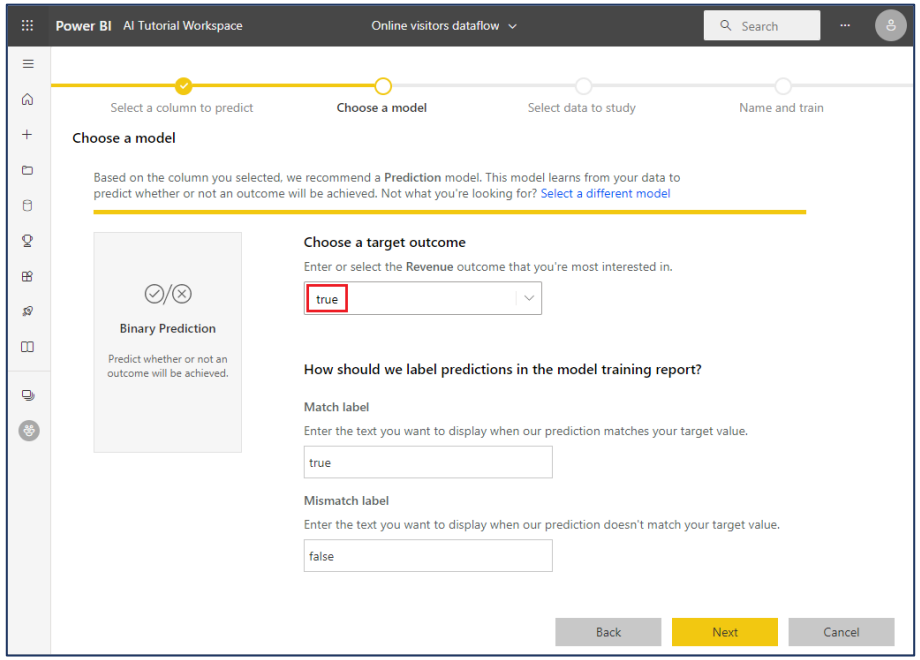


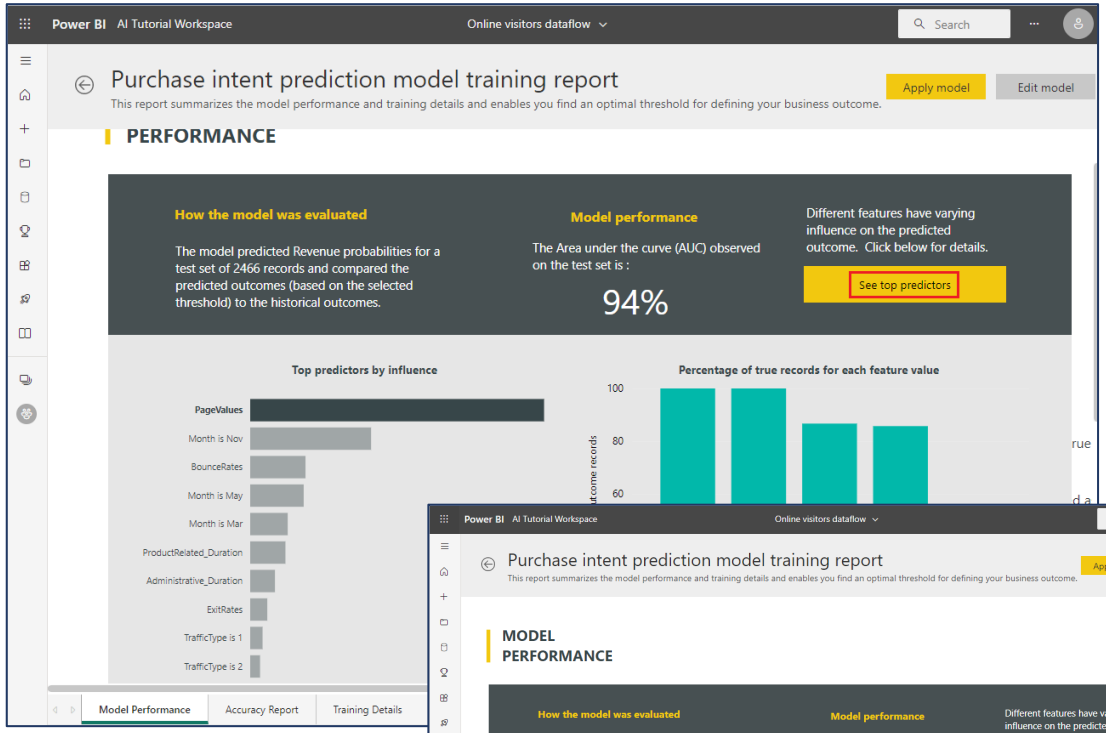
https://learn.microsoft.com/en-us/power-bi/connect-data/service-tutorial-build-machine-learning-model

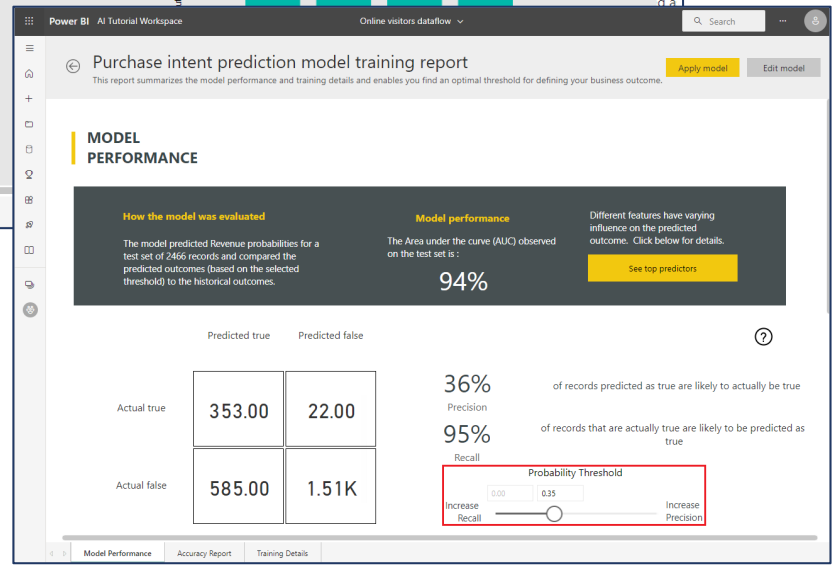
Technology for Data Science as a Service

Example: Power BI







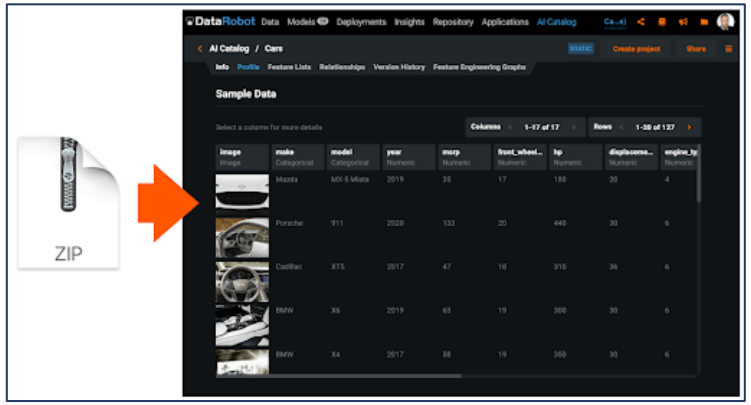


https://learn.microsoft.com/en-us/power-bi/connect-data/service-tutorial-build-machine-learning-model

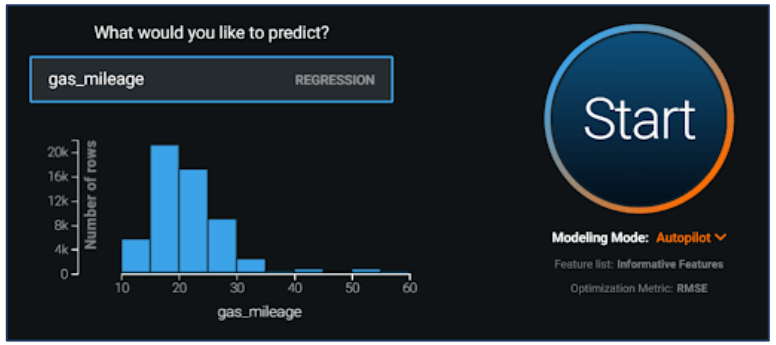
Technology for Data Science as a Service

Example: Visual AI with DataRobot

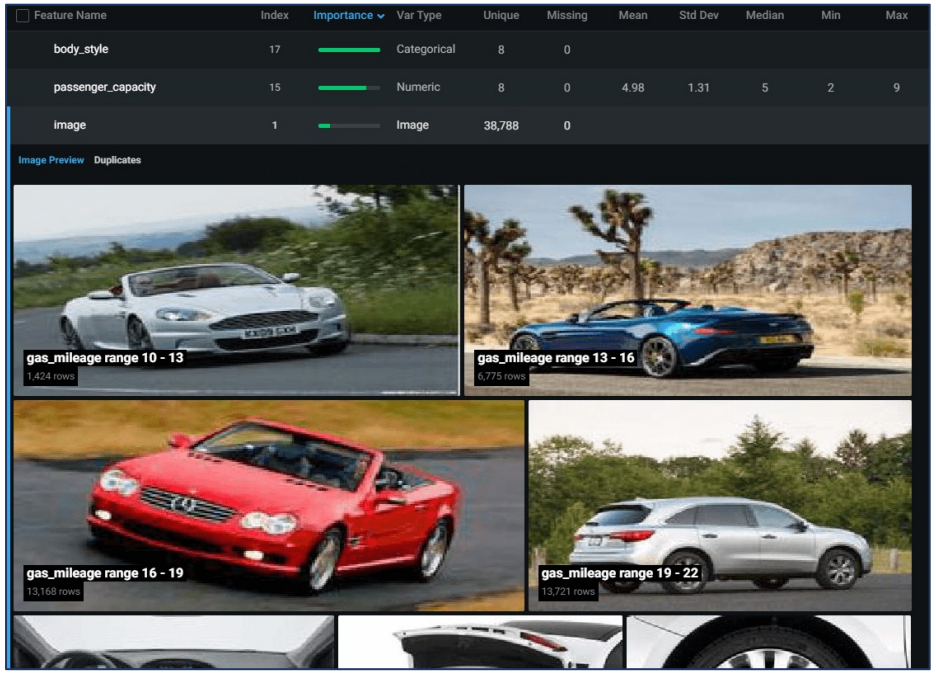
Drag & drop zip file with images for the model



Start building your model



Our automated EDA will show you lots of interesting statistics about your dataset features, including the identification of missing and duplicate images.



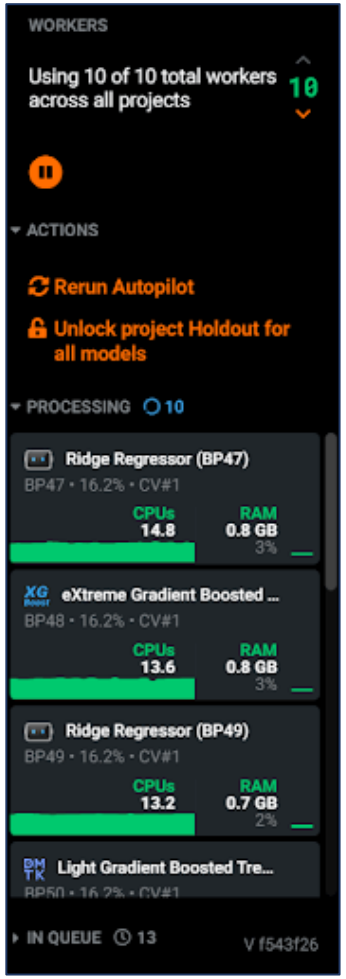
DataRobot will automatically select, train, test and compare a whole host of cutting-edge deep learning algorithms then recommend the best one

https://www.datarobot.com/blog/introducing-visual-ai-for-datarobot-automated-machine-learning/

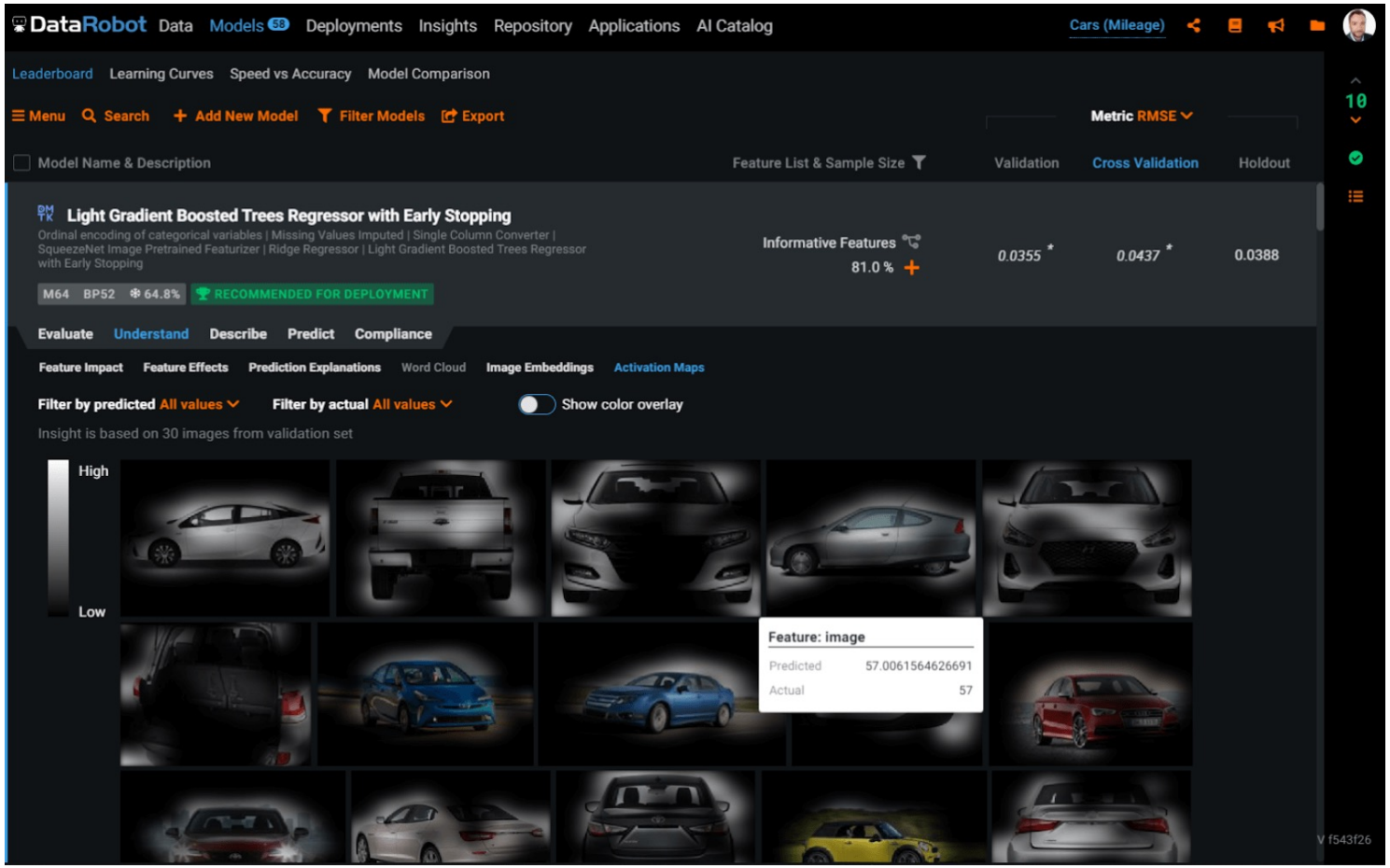
Technology for Data Science as a Service

Example: Visual AI with DataRobot

DataRobot will automatically select, train, test and compare a whole host of cutting-edge deep learning algorithms then recommend the best one



Recommended a model

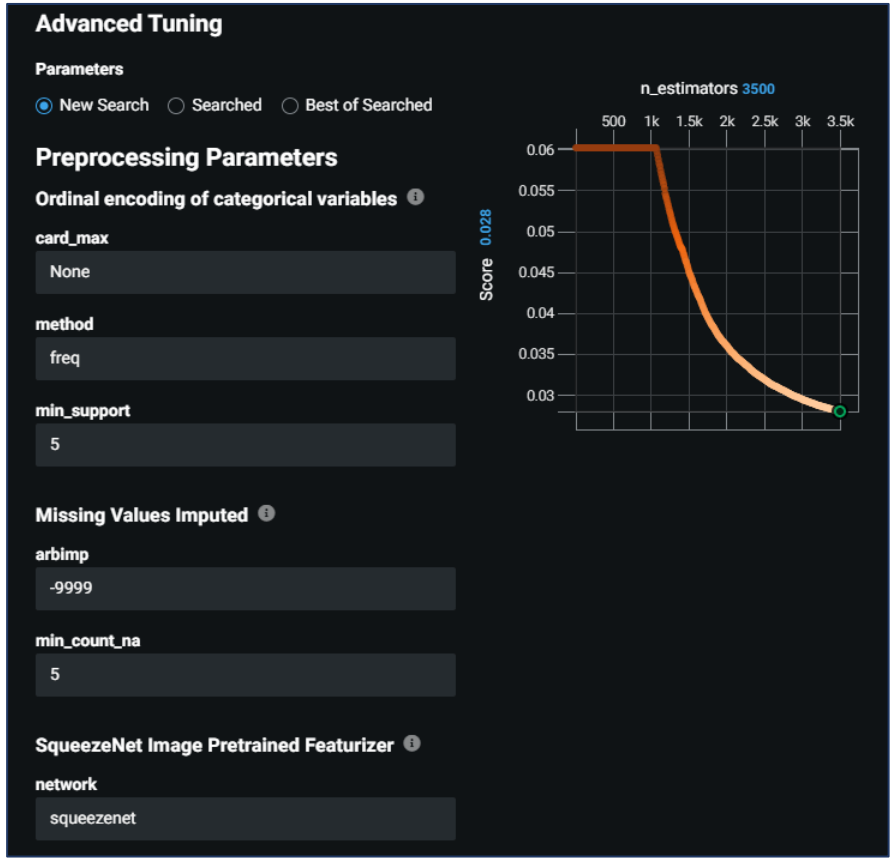


https://www.datarobot.com/blog/introducing-visual-ai-for-datarobot-automated-machine-learning/

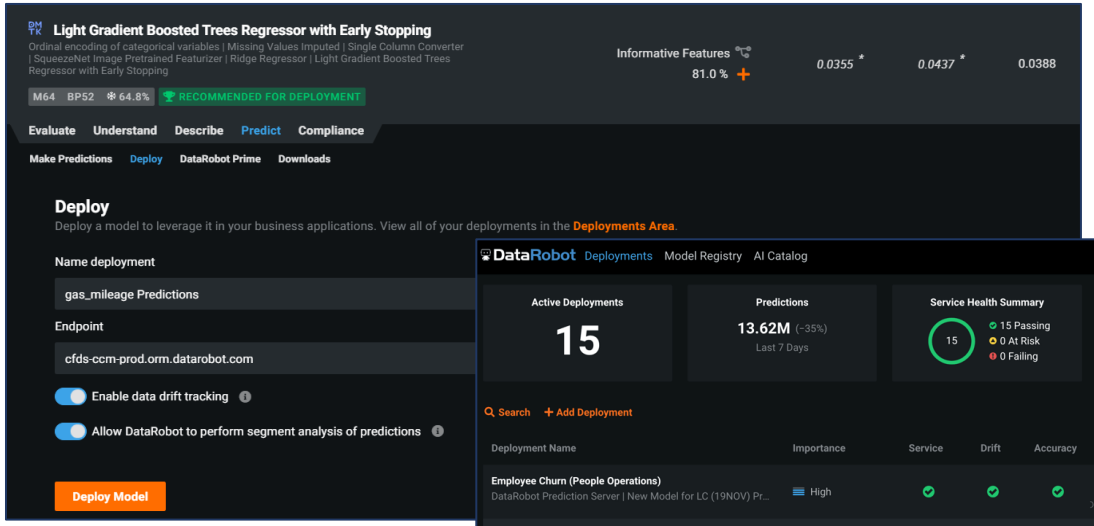
Technology for Data Science as a Service

Example: Visual AI with DataRobot

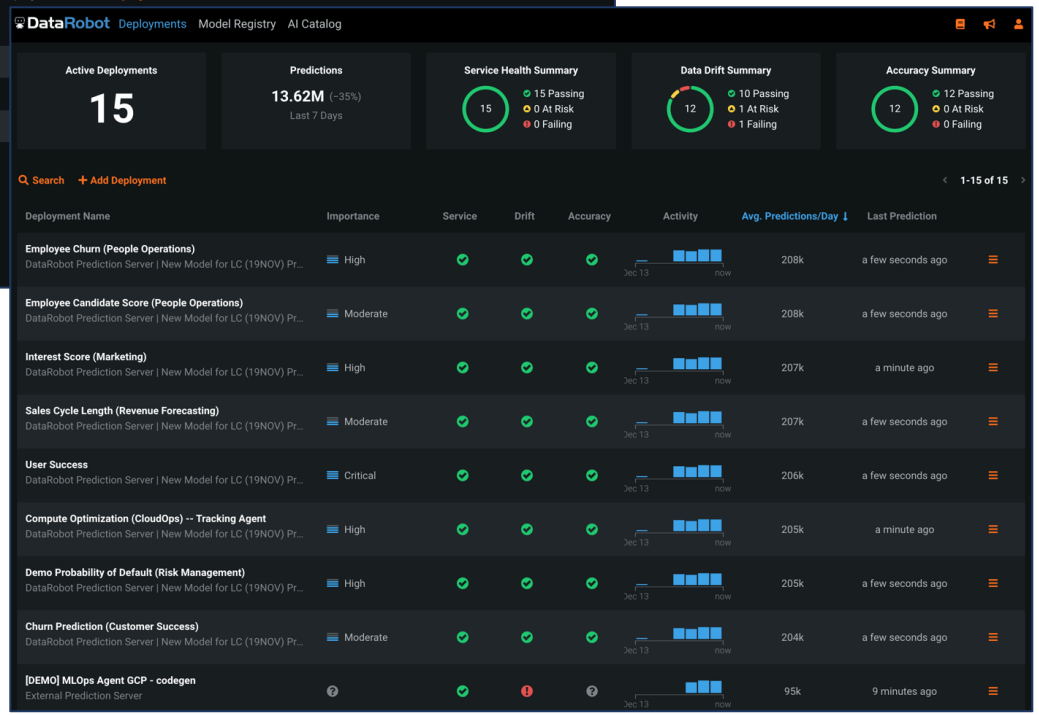
Tune your model



Deploy



Monitor



https://www.datarobot.com/blog/introducing-visual-ai-for-datarobot-automated-machine-learning/

Technology for Data Science as a Service

Data Science as a Service

- Ingestion of data
- EDA – Exploratory Data Analysis
- Feature Engineering
- Data Science projects such as Chatbots, Computer vision, Fraud Detection and more

There are many tools that can be used for this purpose

- **Dataiku** - collaborartive and user friendly
- **Knime** – open-source focused on business inteliggence
- **H2O.ai** – open-source focused on EDA and auto-ml

https://scalefresh.com/5-awesome-tools-for-citizen-data-scientists/

Data Science as a Service

Takeaways

- Data Science gives competitive advantage to organizations **however Citizen Data Scientists are not always empowered**

- Citizen Data Scientists produce more value when closely advised, or working, with Expert Data Scientists

- Provide opportunities for more people contribute to the data analytics industry by supporting learning skills for
  - **Data Literacy & Story Telling**
  - **Machine Learning Literacy**
  - **Technology for Data Science as a Service**

Thank you!

Filipa Peleja

https://www.linkedin.com/in/peleja/

Lead Data Science Engineer @ LEGO Group
Lead Data Science Instructor @ CodeOp

CodeOp *– Bridging technical talent development with DE&I through our* ***international tech school for women+***

Image from "but Asking Can Hurt," She Writes - Ask Questions