

Data Science в Microsoft Azure: как создавать модели машинного обучения и не думать об инструментах и ресурсах

Анна Фенюшина
Архитектор
Softline



Содержание

Что такое Искусственный Интеллект и Машинное Обучение
Сценарии использования, практическое применение

Преимущества использования Azure для DataScience

Краткий обзор сервисов для Машинного Обучения и
Искусственного Интеллекта в Azure для разных категорий
специалистов

Сервисы и ресурсы для Data Science:

- Виртуальные машины для Data Science;
- Инструменты для Data Scientists;
- Оптимизация процесса создания моделей

Варианты развертывания моделей в Azure

Вопросы и ответы

Большие данные, продвинутая аналитика и искусственный интеллект ускоряют цифровую трансформацию, помогая находить новые закономерности и принимать взвешенные решения на базе полученных знаний

Искусственный Интеллект во всех бизнес процессах

Горизонтальные

Маркетинг
IT helpdesk
Финансовые прогнозы
Управление талантами
Клиентский сервис
Рекомендации
Скоринг



Банки



Медицина



Торговля



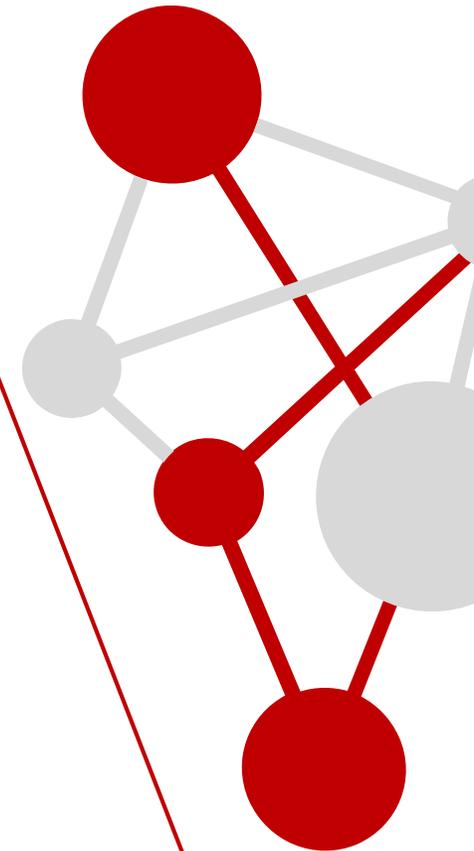
Телеком



Производство

Вертикальные

Оптимизация инфраструктуры
Упреждающее обслуживание
Инвентаризация
Логистика
Анализ изображений
Динамическое ценообразование
Управление рисками
Контроль качества



Краткий обзор сервисов для Машинного Обучения

Сервис Чат-ботов

Ускоренная разработка чат-ботов с использованием ИИ.

Готовые Когнитивные Сервисы

Готовые сервисы по распознаванию образов, речи, пониманию языка и интеллектуальному поиску.

Сервис Машинного Обучения

Сервис по созданию экспериментов и управлению развертыванием моделей Машинного Обучения.



ГОТОВЫЕ КОГНИТИВНЫЕ СЕРВИСЫ



Vision

Video Indexer
Computer Vision
Face
Emotion
Content Moderator
Custom Vision



Speech

Speaker Recognition
Bing Speech
Custom Speech
Translator Speech
Unified Speech
Speech to Text w. Custom Speech
Text to Speech w. Custom Voice
Speech Translation w. Custom Translator



Language

Text Analytics
Bing Spell Check
Translator Text
Language Understanding
QnA Maker



Web Search

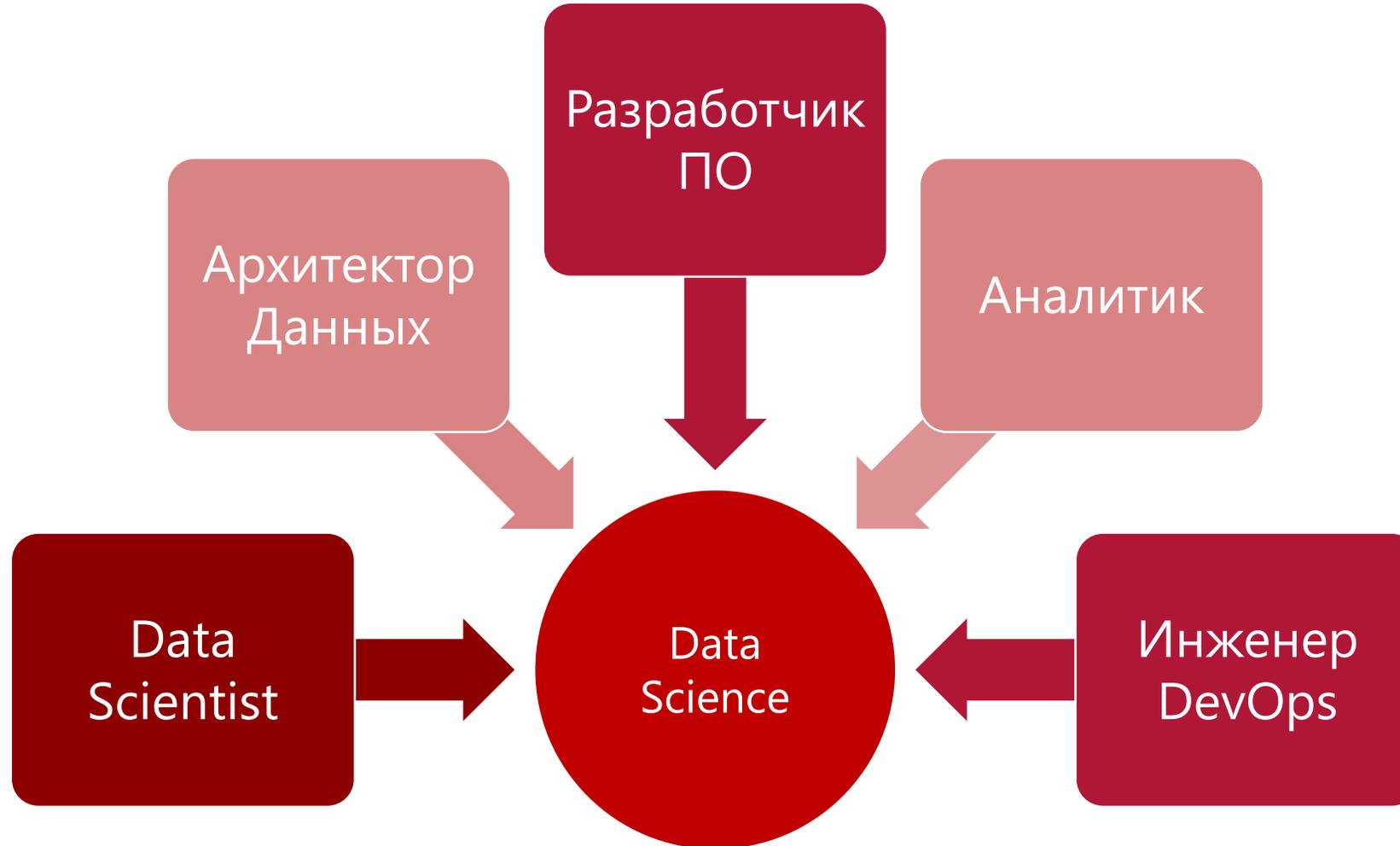
Bing Entity Search
Bing Autosuggest
Bing Search
Web Search
Image Search
News Search
Video Search
Bing Visual Search
Bing Custom Search
Bing Local Business Search



Labs

Project Gesture
Project Local Insights
Project Academic Knowledge
Project Entity Linking
Project Knowledge Exploration
Project Event Tracking
Project Answer Search
Project URL Preview
Project Anomaly Finder
Project Conversation Learner
Project Personality Chat

Категории специалистов, работающих с Искусственным Интеллектом



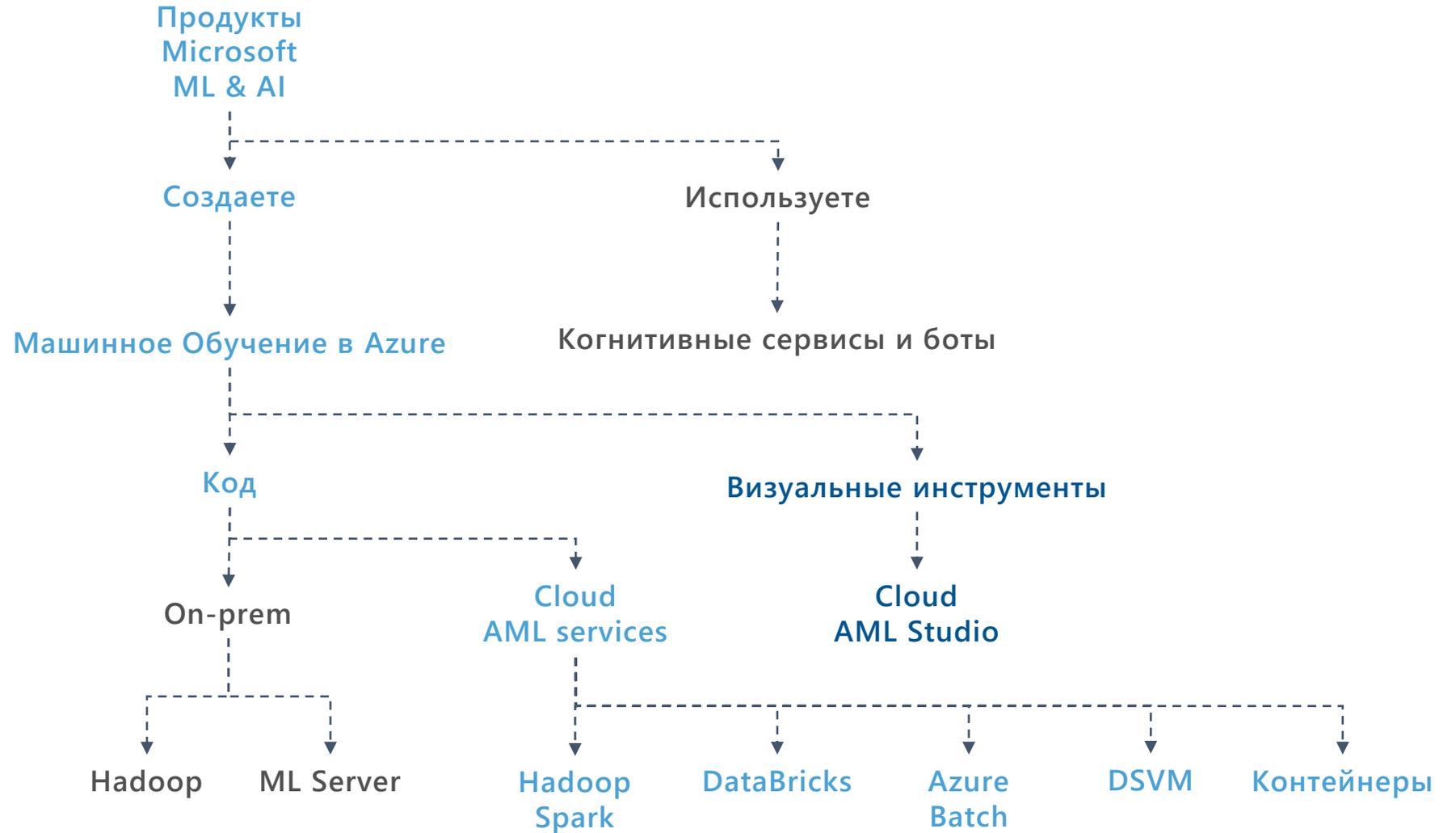
Выбор сервисов Искусственного Интеллекта

Создаете свои модели или используете готовые сервисы?

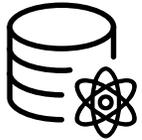
Какие инструменты Вы хотите использовать?

Где Вы будете разворачивать?

Какой движок Вы хотите использовать?



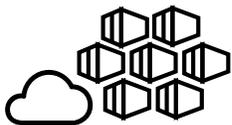
Преимущества использования Azure для DataScience



Увеличивайте продуктивность DataScientists



Улучшайте качество моделей



Развертывайте и управляйте своими моделями в любых средах

Разработано с учетом опыта Data Science

Виртуальные машины с GPU

Масштабируемость и низкие задержки

Интеграция с популярными Python IDEs

Ролевая модель доступа к управлению сервисами и ресурсами (Role-based access controls)

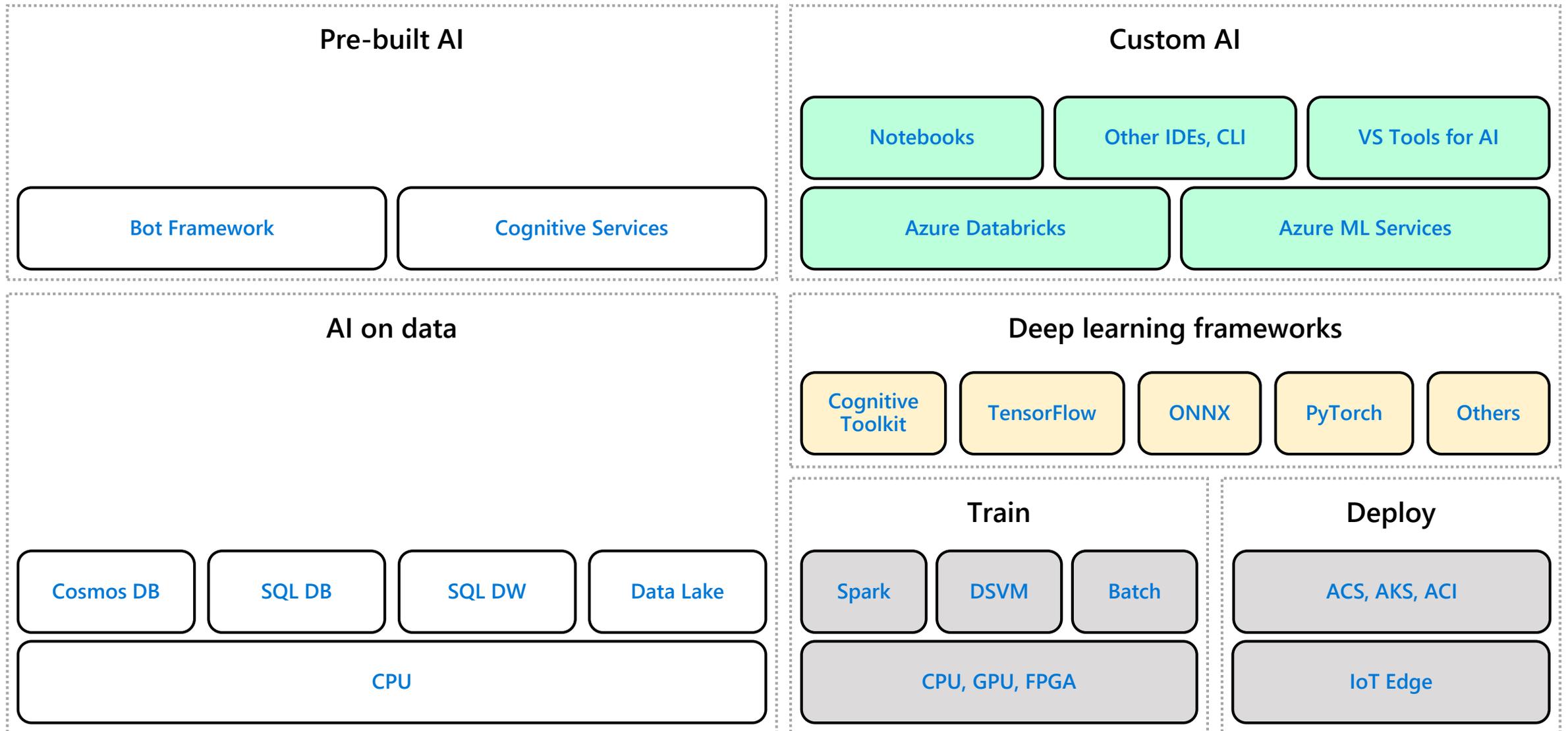
Контроль версий моделей, наборов данных, экспериментов

Автоматизация создания моделей



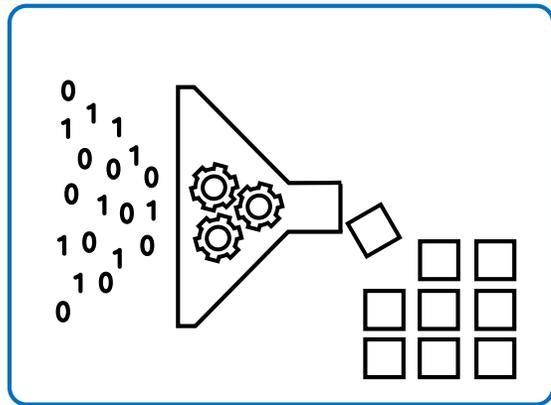
Бесшовная интеграция с другими сервисами Azure

Сервисы Azure, используемые для DataScience

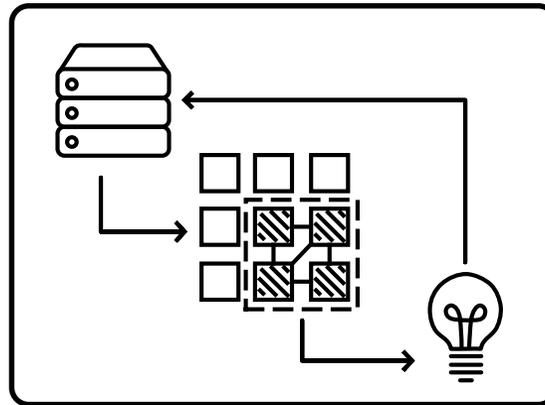


Этап создания модели

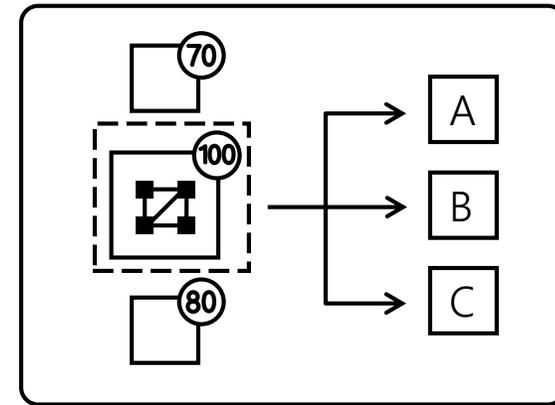
Машинное обучение



Подготовка данных и модели

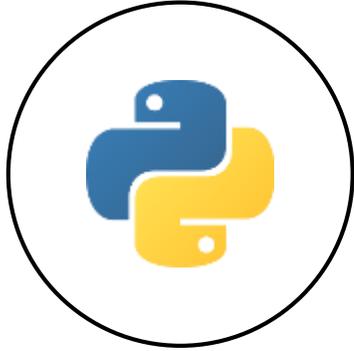


Обучение и оценка модели

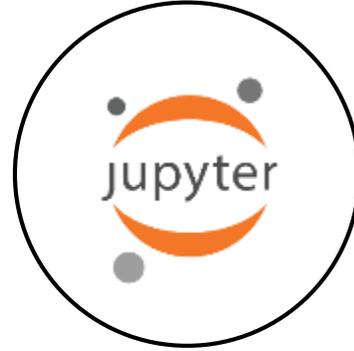


Развертывание и управление

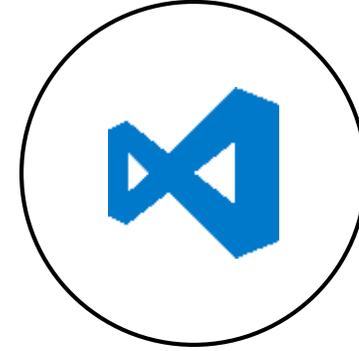
Подготовка данных и создание моделей



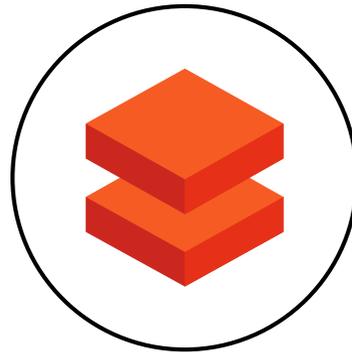
PyCharm



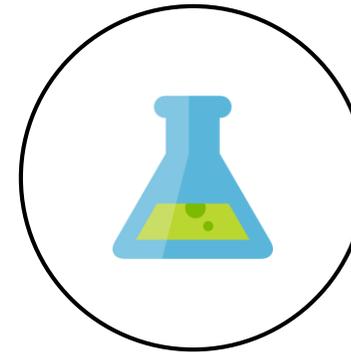
Azure и Jupyter Notebooks



Visual studio tools for AI



DataBricks Notebooks



ML Studio

Популярные Frameworks и Azure Machine Learning Python SDK

Use your favorite deep learning frameworks



TensorFlow



PyTorch



Scikit-Learn



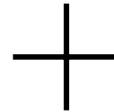
MXNet



Chainer



Keras



Without getting locked into one framework



Community project created by Facebook and Microsoft

Use the best tool for the job. Train in one framework and transfer to another for inference



Azure Machine Learning Python SDK

Инструменты Visual Studio для Машинного Обучения

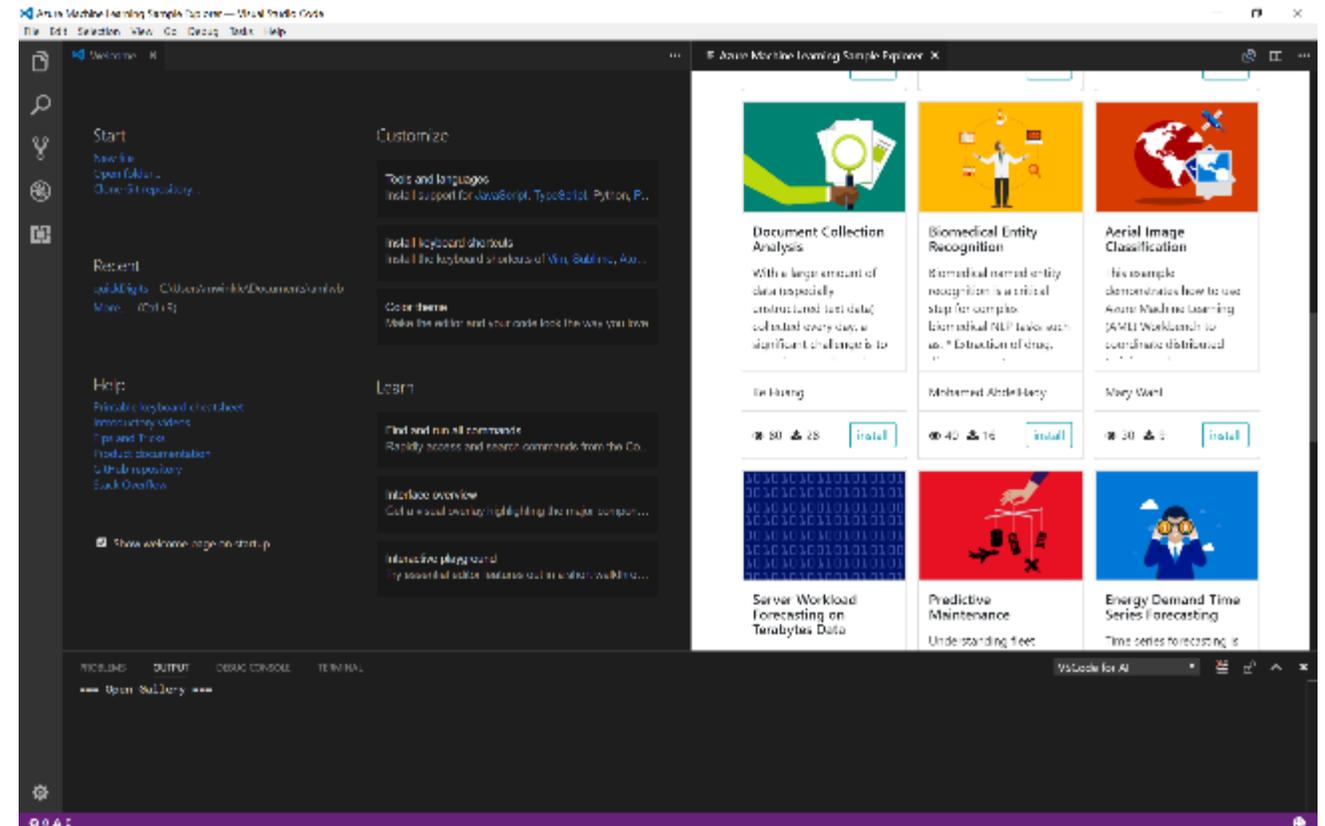
Дополнение к Visual Studio Code для интеграции с Azure ML

Среда разработки от нового проекта до развертывания моделей

Поддержка обучения на удаленных ресурсах

Управление задачами

Всё лучшее от Visual Studio, Python, Jupyter, Git и т.д.

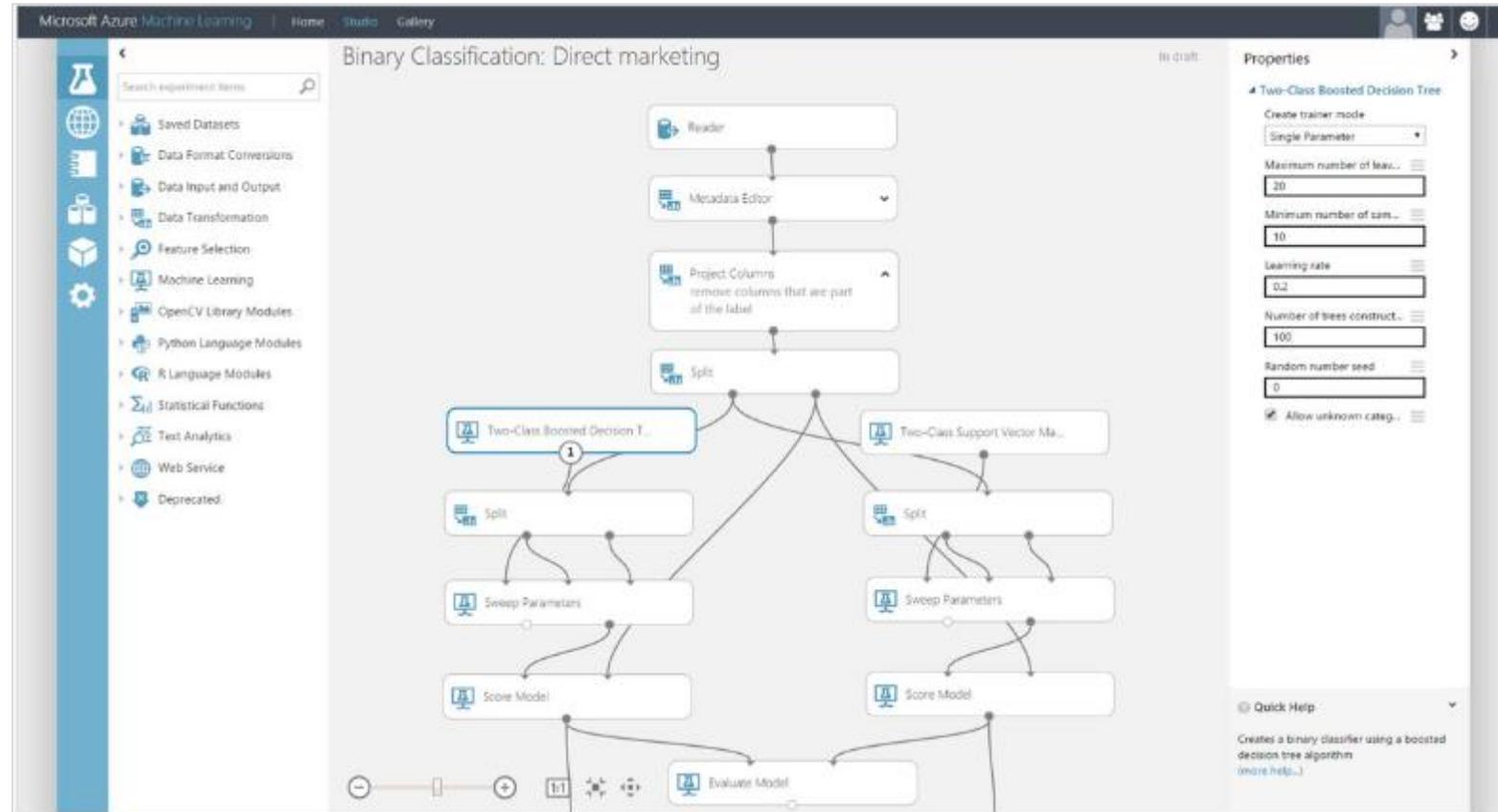


Визуальный интерфейс (отличия от Azure Machine Learning Studio)

Графический инструмент создания, обучения, оценки и развертывания моделей.

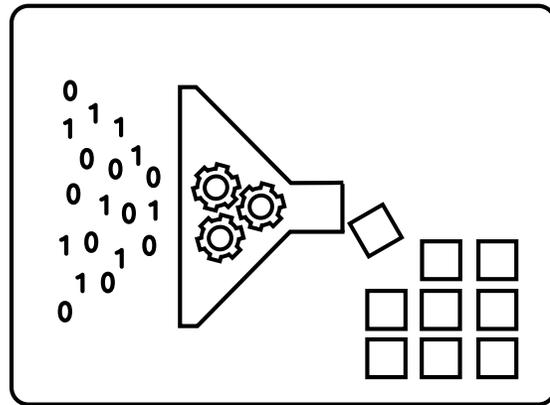
Очень похож на Azure ML Studio.

- Быстрое создание экспериментов
- Готовые модули для подготовки данных, тренировки и оценки модели
- (отличие) Возможность выбрать среду обучения. Развертывание в AKS

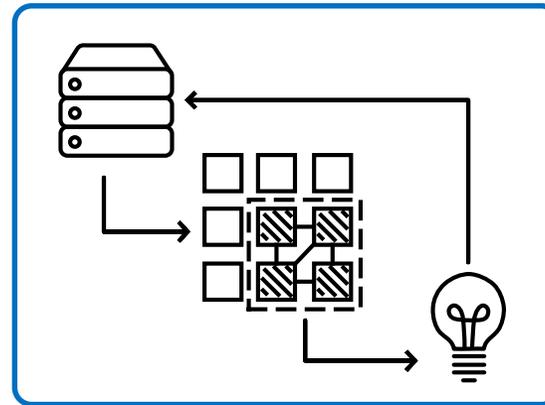


Этап обучения и оценки модели

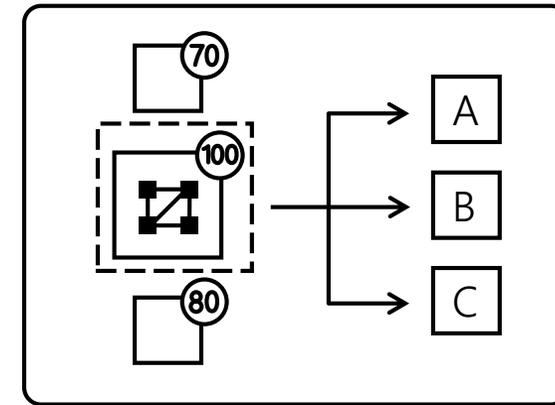
Машинное обучение



Подготовка данных и модели



Обучение и оценка модели

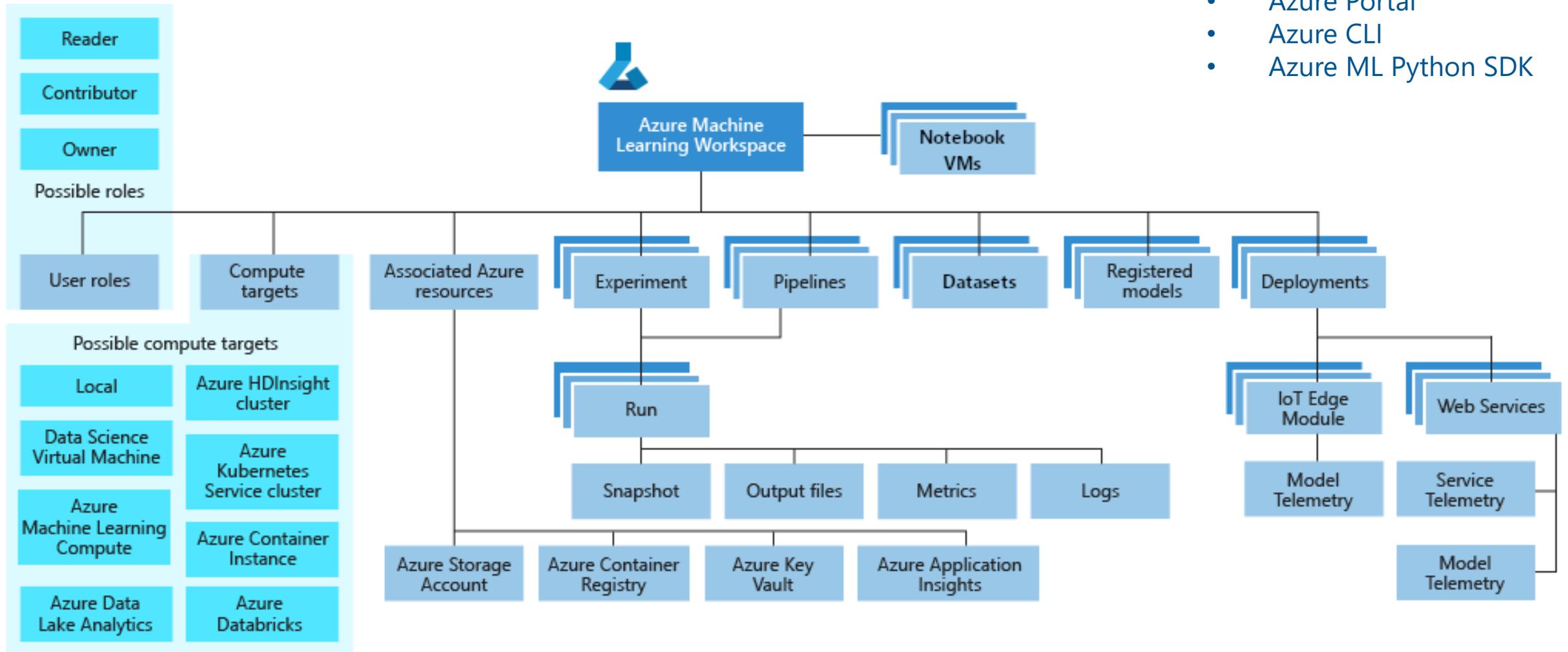


Развертывание и управление

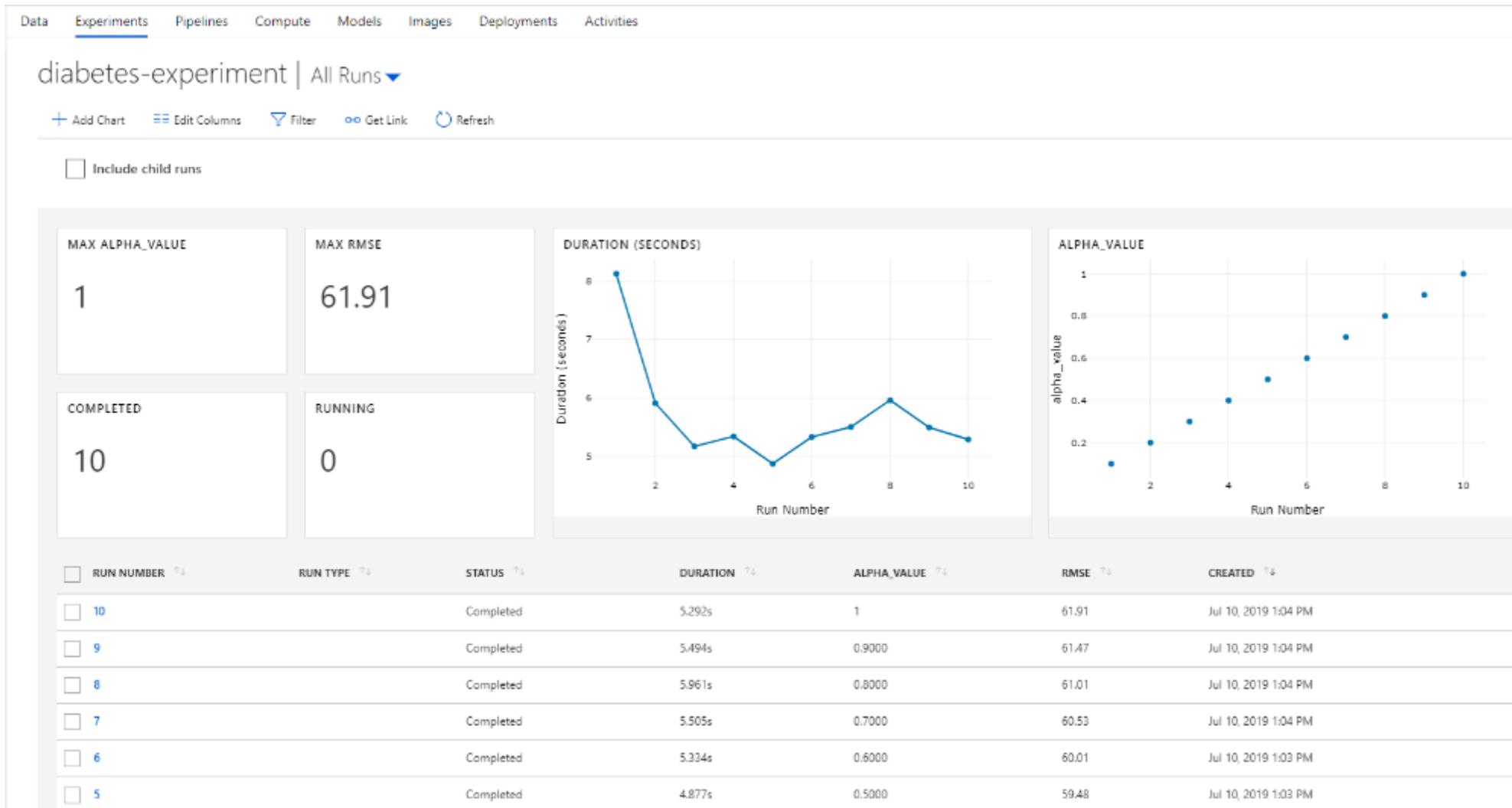
Служба Azure Machine Learning

Управление:

- Azure Portal
- Azure CLI
- Azure ML Python SDK



Оптимизация процесса создания моделей и выбора наилучшей



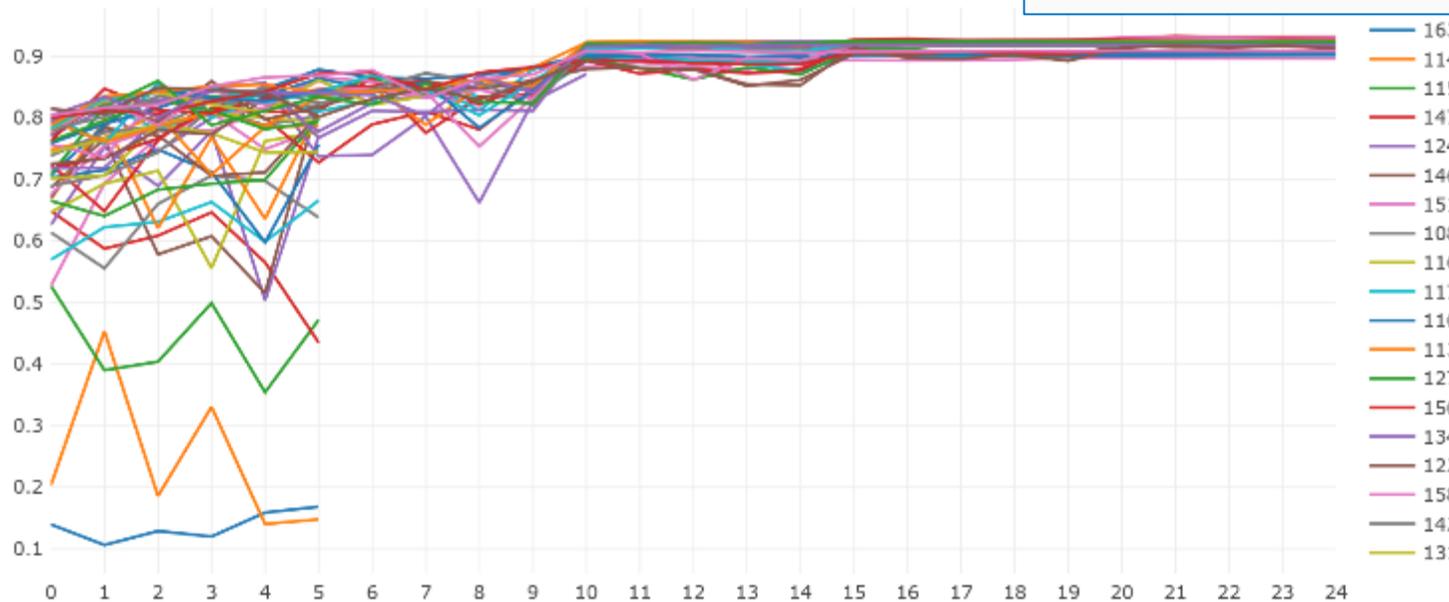
Tune HyperParameters и другие Estimators (scikit-learn, PyTorch, TensorFlow, Chainer and Keras)

```
from azureml.train.hyperdrive import HyperDriveConfig
hyperdrive_run_config = HyperDriveConfig(estimator=estimator,
                                         hyperparameter_sampling=param_sampling,
                                         policy=early_termination_policy,
                                         primary_metric_name="accuracy",
                                         primary_metric_goal=PrimaryMetricGoal.MAXIMIZE,
                                         max_total_runs=100,
                                         max_concurrent_runs=4)
```

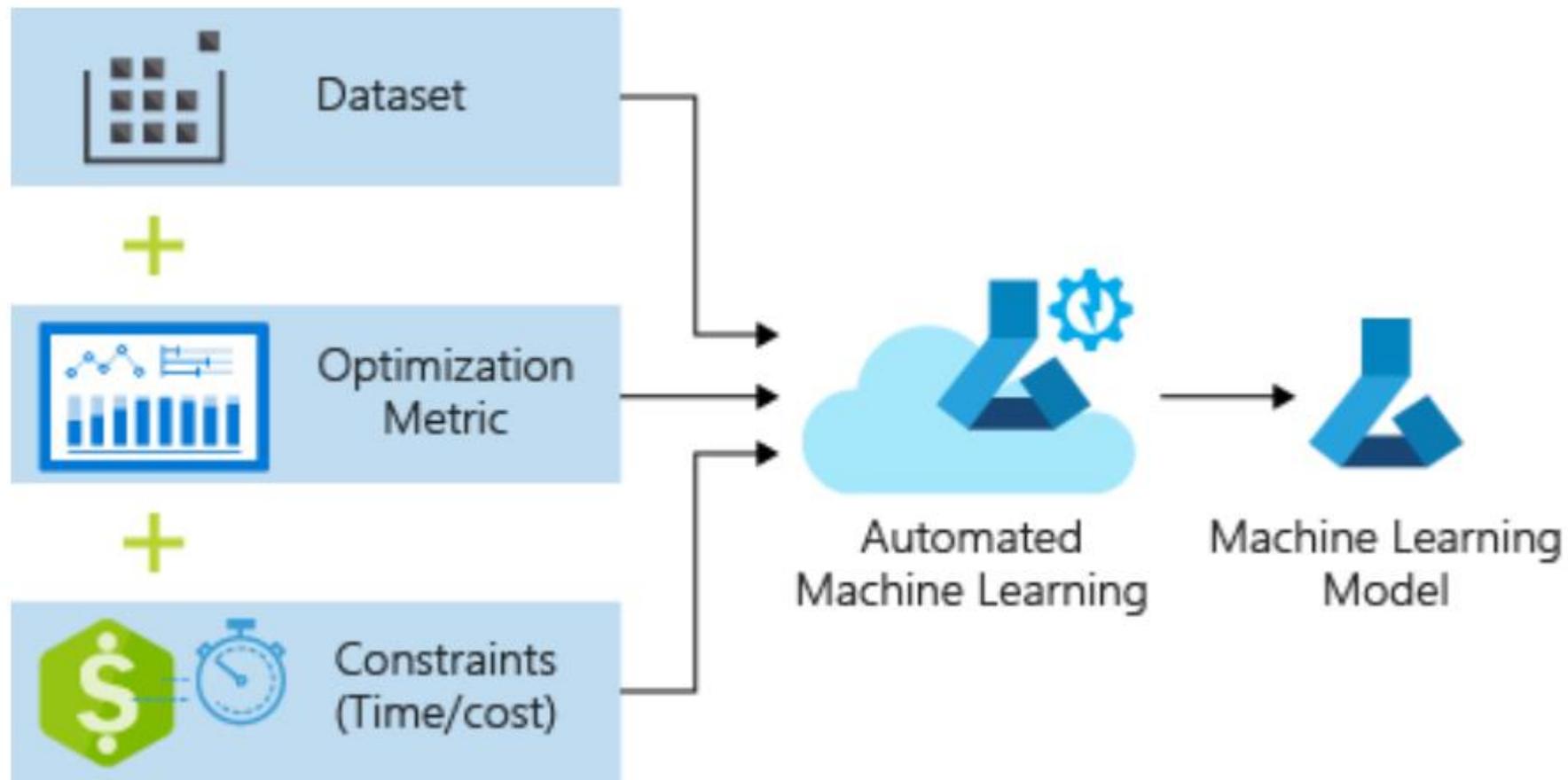
```
from azureml.core.experiment import Experiment
experiment = Experiment(workspace, experiment_name)
hyperdrive_run = experiment.submit(hyperdrive_run_config)
```

```
from azureml.widgets import RunDetails
RunDetails(hyperdrive_run).show()
```

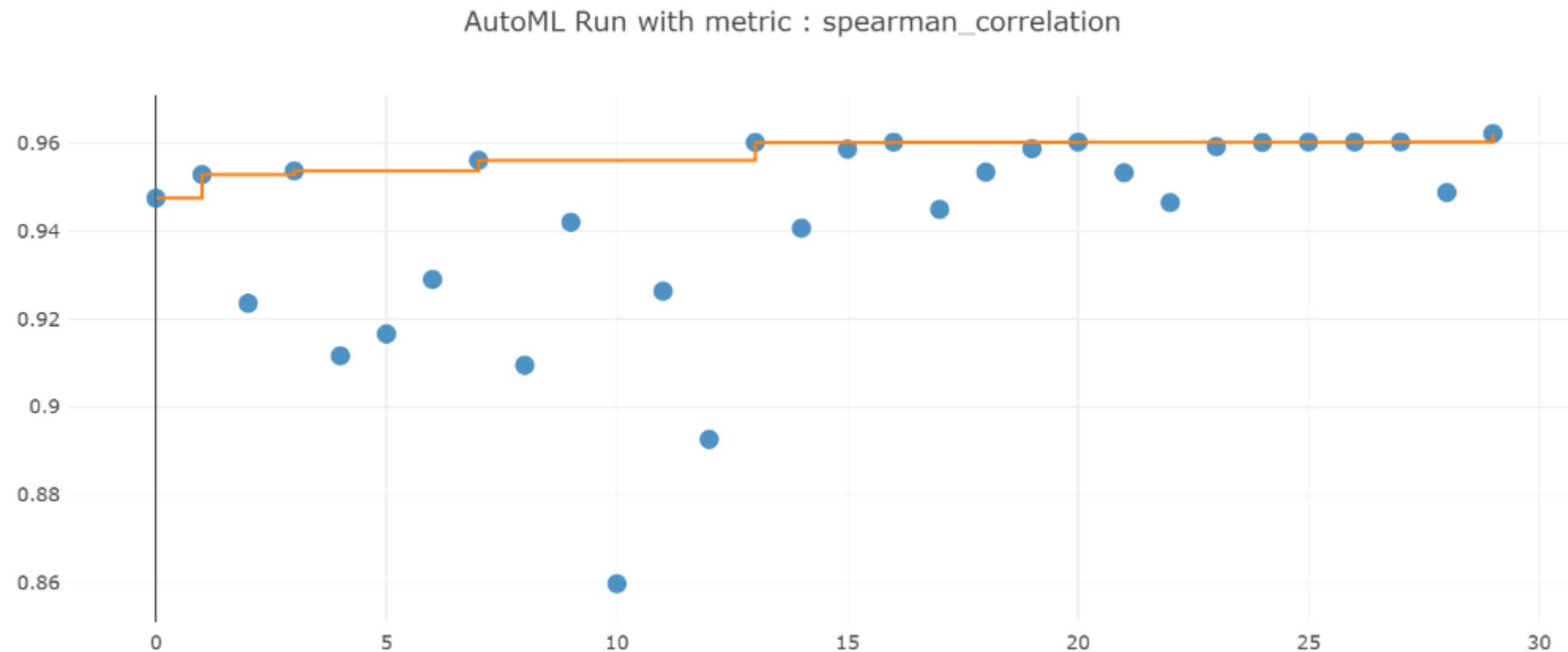
HyperDrive Run Primary Metric : accuracy



Автоматизация создания моделей



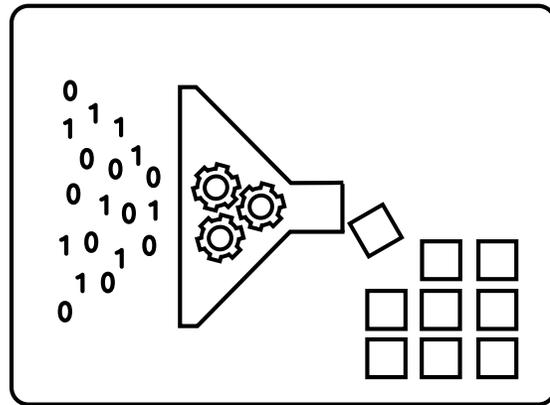
Автоматизация создания моделей



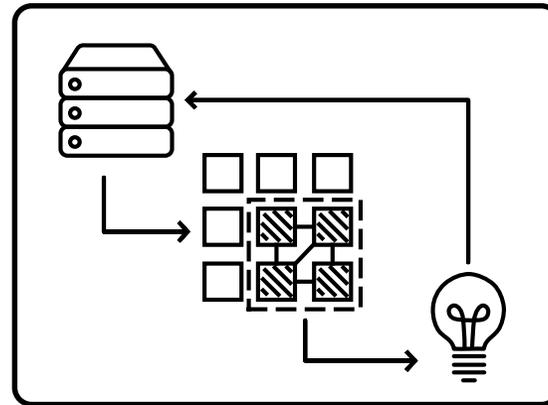
[Click here to see the run in Azure portal](#)

Этап развертывания модели

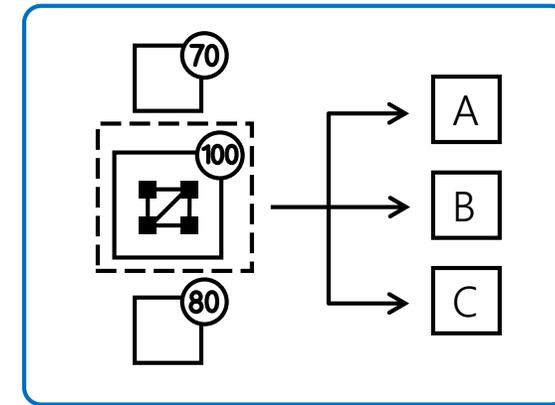
Машинное обучение



Подготовка данных и модели



Обучение и оценка модели

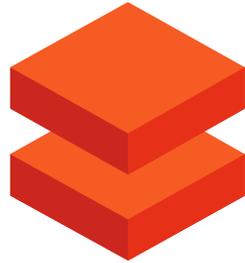


Развертывание и управление

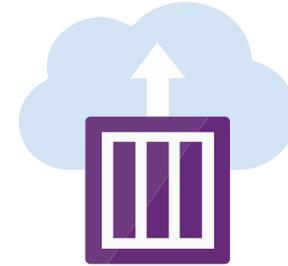
Ресурсы для обучения и развертывания моделей



Локальные
рабочие
станции



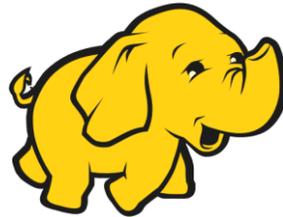
Кластер
DataBricks



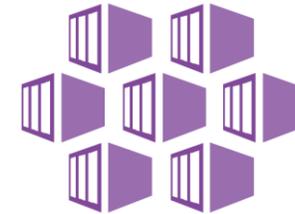
Контейнеры
(Container
Instances)



Azure ML
Compute



HDInsight
Spark



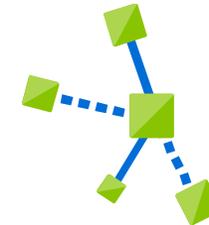
Azure
Kubernetes
(AKS)



Виртуальные
машины для
DataScience



Azure
Batch



IoT Edge

Виртуальные машины для DataScience



Быстро, практически без дополнительной настройки запускайте ресурсы для обучения и оценки моделей



Созданы и настроены специально для использования с GPU



Масштабируйте ресурсы по требованию как вертикально, так и горизонтально

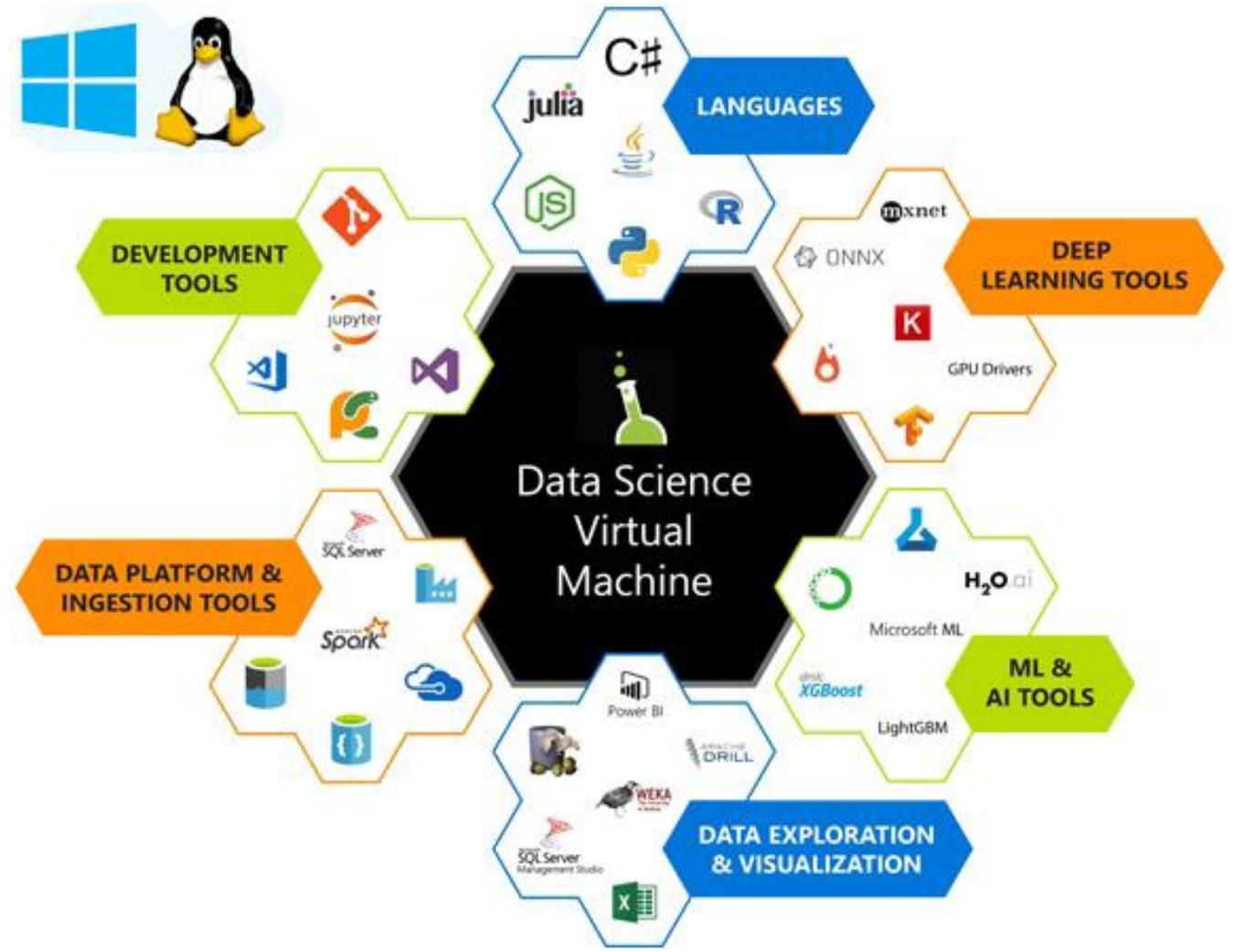
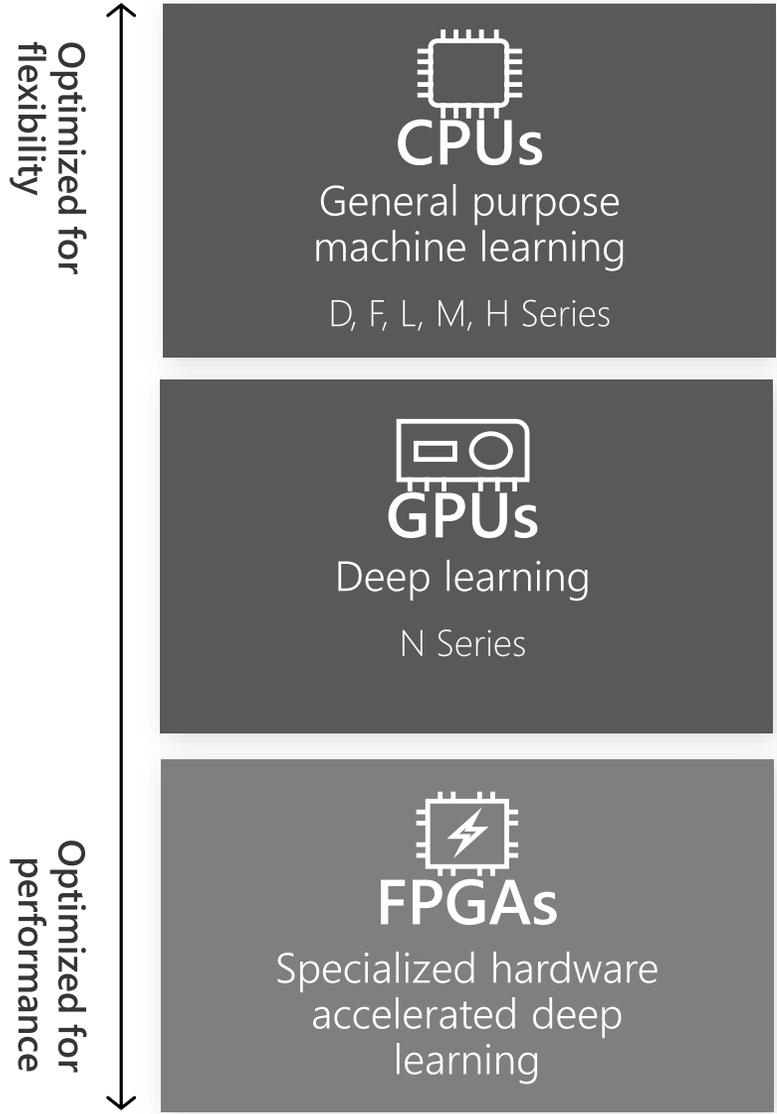


Тесная интеграция с Сервисом Машинного Обучения

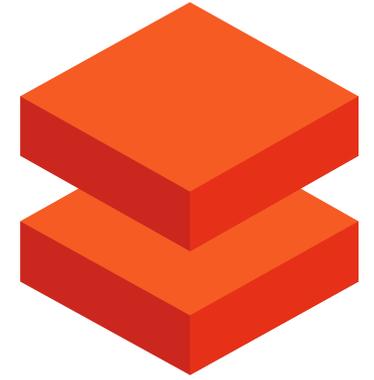


Быстрый старт с предустановленными примерами, темплейтами, готовыми ноутбуками и наборами данных

Виртуальные машины для DataScience



Data Bricks



Быстрая и простая аналитическая платформа для совместной работы на основе Apache® Spark™, оптимизированная для Azure



Best of Databricks

Best of Microsoft



Разработано в сотрудничестве с основателями Apache Spark



Настройка в один клик; оптимизированные рабочие процессы, простое планирование заданий, автоматическое масштабирование



Интерактивное рабочее пространство, обеспечивающее сотрудничество между учеными, инженерами и бизнес-аналитиками. Ролевое управление доступом



Встроенная интеграция со службами Azure (Power BI, SQL DW, Cosmos DB, Blob Storage). Мощные ноутбуки



Безопасность Azure корпоративного уровня (интеграция с Active Directory, соответствие требованиям, SLA корпоративного уровня)

Batch AI, Edge, Containers & AKS



Azure Container Instances



Запускайте контейнеры без необходимости администрировать сервера, виртуальные машины или дополнительные сервисы



Публикуйте свои сервисы для использования



Контейнеры выполняются в изолированных средах, что повышает безопасность

Контейнеры



Azure Kubernetes Service



Создайте кластер Kubernetes за несколько минут всего несколькими кликами мыши



Используйте привычные инструменты Kubernetes, такие как Helm и Draft



Быстро и эффективно масштабируйте свои ресурсы без остановки работы сервиса



Batch



Автоматически масштабируйте кластера с GPU или CPU



Создавайте модели с длительной пакетной обработкой данных, последовательными экспериментами и обучением



Поддержка любых framework-ов машинного обучения



Edge



Создавайте умные устройства



Сокращайте время реакции на события

В качестве заключения

Готовые когнитивные сервисы

Для упрощения создания решений с ИИ



Vision



Speech



Language



Search

Популярные frameworks

Для создания комплексных решений глубокого обучения



Pytorch



TensorFlow



Keras



Onnx

Высокопроизводительные сервисы

Для команд разработчиков и data science



Azure Databricks



Azure Machine Learning



Machine Learning VMs

Мощная инфраструктура

Для ускорения глубокого обучения



CPU



GPU



FPGA

Гибкое развертывание

Развертывание и управление моделями как в облаке, так и на устройствах Интернета Вещей



On-premises



Cloud



Edge

CSP – самый выгодный канал лицензирования для бизнеса любого уровня.



Свободное масштабирование ресурсов по потребностям

Нет обязательств по объему потребления (min/max)

Вы платите только за те ресурсы, которые потребили

Использование ресурсов для себя / аффилиатов / сторонних компаний

Почему Softline: мы делаем управление подписками Azure простым и удобным

Единая платформа управления подписками CSP: Active Platform от Softline

Что вы можете с личным кабинетом Active Platform:

- Пополнять баланс и получать счета за потребление
- Получать триальные версии продуктов (применимо к Office 365)
- Добавлять или убирать подписки
- Видеть данные по объему потребления Azure в актуальных ценах (эта опция не доступна в учетной системе Microsoft)
- Видеть все ваши подписки CSP на Office 365 и Azure в одном кабинете
- Детализировать информацию по потребляемым ресурсам, группам ресурсов, тегам.

Вы самостоятельно управляете своими Балансами и объемами потребления



GO GLOBAL



GO CLOUD



GO INNOVATIVE