Az COMMUNITY

Conference 2022

Asia's Largest Azure Community Conference

#AzConfDev

Praveen Raghuvanshi

Cloud Architect, Harman International

#AzConfDev

Applying Data Science and Machine Learning Using Microsoft .Net

Praveen Raghuvanshi

Acknowledgements

- Dataset – John Hopkins University CSSE
- Python Visualization – Akshay Sb
- Organizers – AzConf 2022
- Sponsors

Identify Language

Maidin mhaith

Guten Morgen

おはよ ございます

शुभ प्रभात

Good Morning

Programming Languages

Source: https://insights.dice.com/2019/10/24/tiobe-top-eight-top-ten-open/

Data Science – Exploratory Data Analysis(EDA)

As per Wiki…

"In statistics, exploratory data analysis (EDA) is an approach of analyzing data sets to summarize their main characteristics, often with visual methods."

- Buy a House
  - No of bedrooms
  - Type (Apartment, Villa)
  - Locality
  - Crime Rate

Source: https://insights.dice.com/2019/10/24/tiobe-top-eight-top-ten-open/

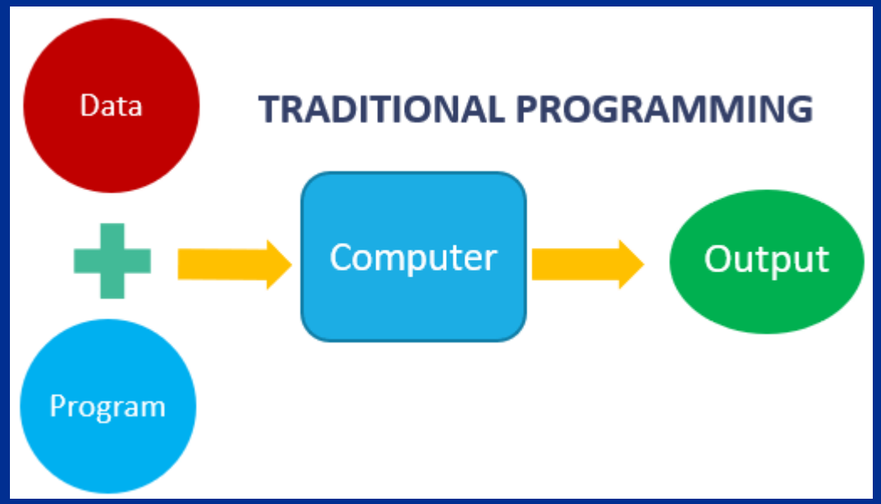
EDA Datastructure : DataFrame

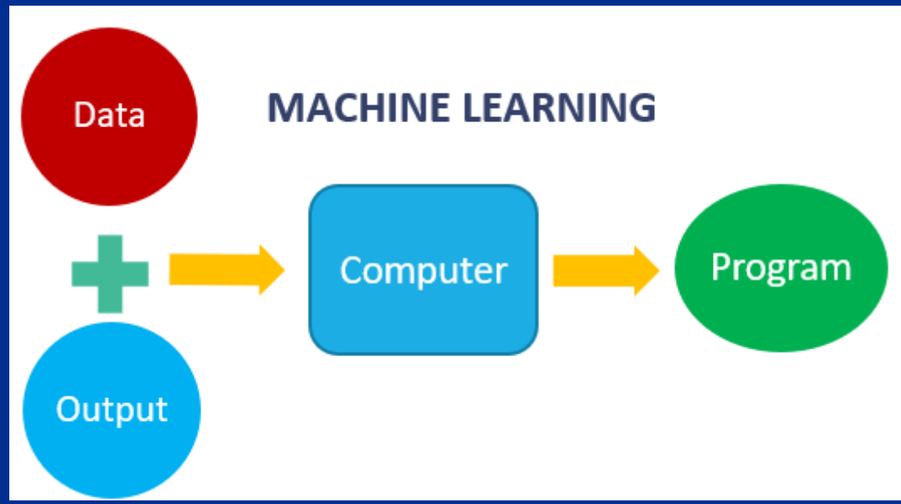
	NAME	AGE	DESIGNATION
1	a	20	VP
2	b	27	CEO
3	c	35	CFO
4	d	55	VP
5	e	18	VP
6	f	21	CEO
7	g	35	MD

	age	designation	name
0	20	VP	a
1	27	CEO	b
2	35	CFO	c
3	55	VP	d
4	18	VP	e
5	21	CEO	f
6	35	MD	g

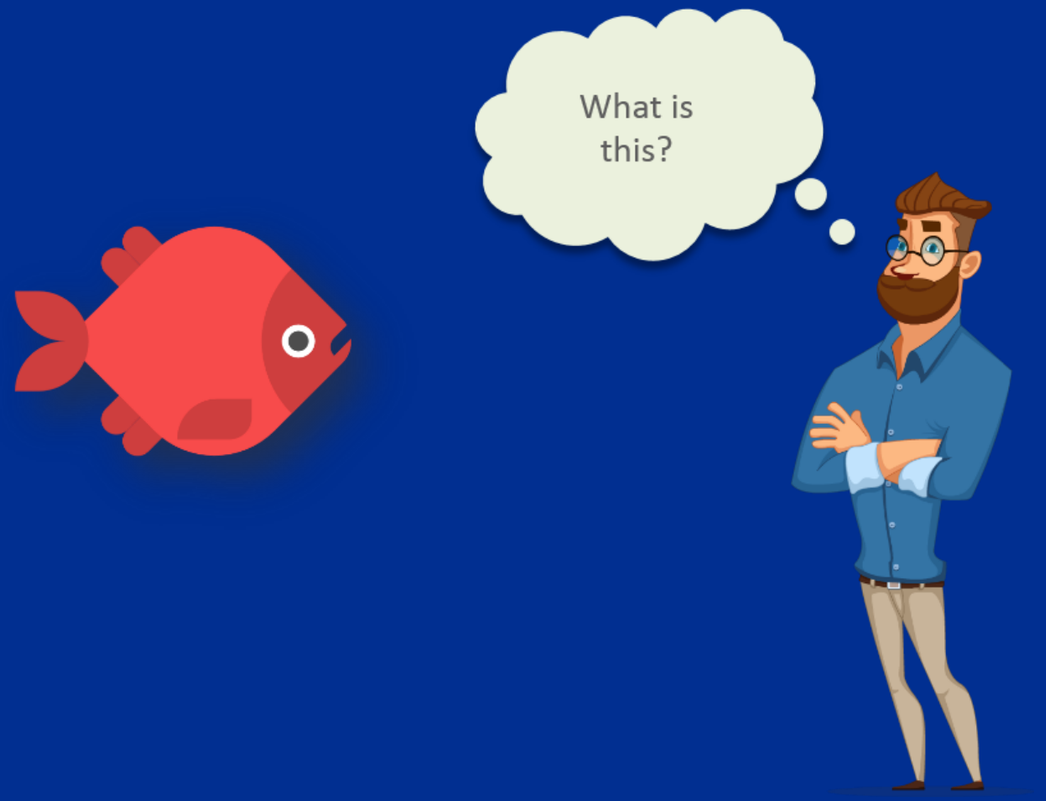
Machine Learning

- Machine learning is getting computers to program themselves
- If programming is automation, then machine learning is automating the process of automation



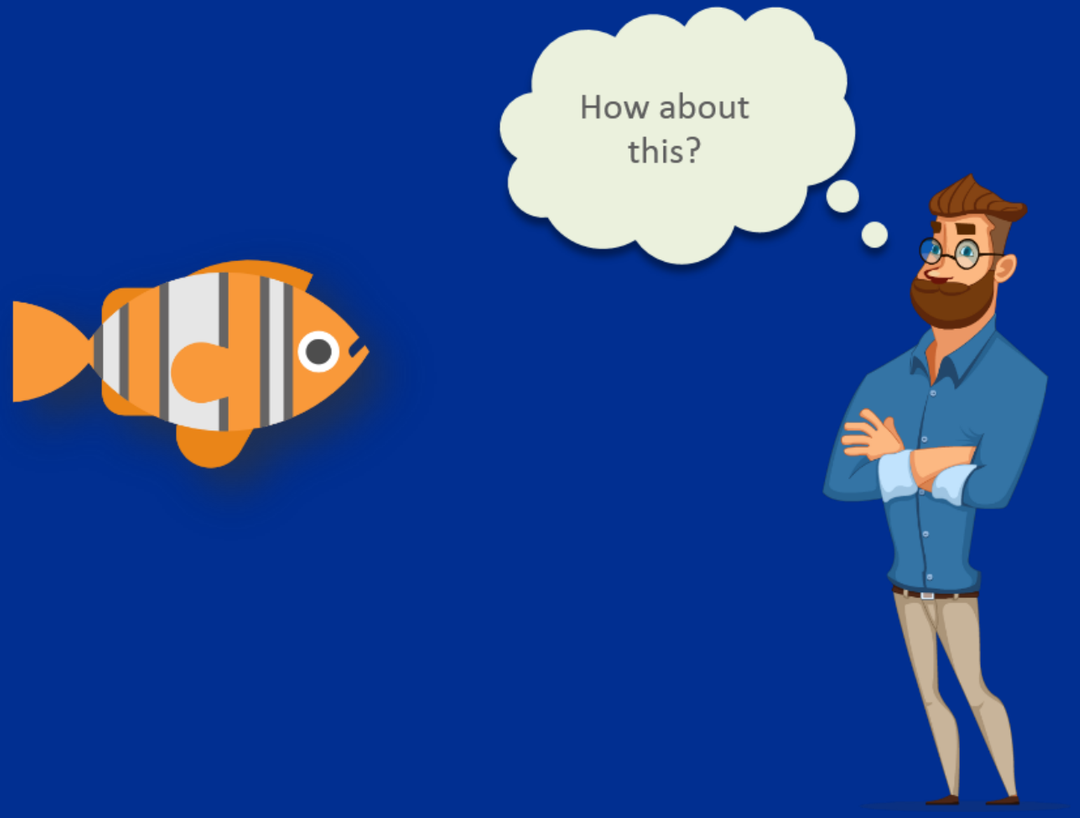


Machine Learning



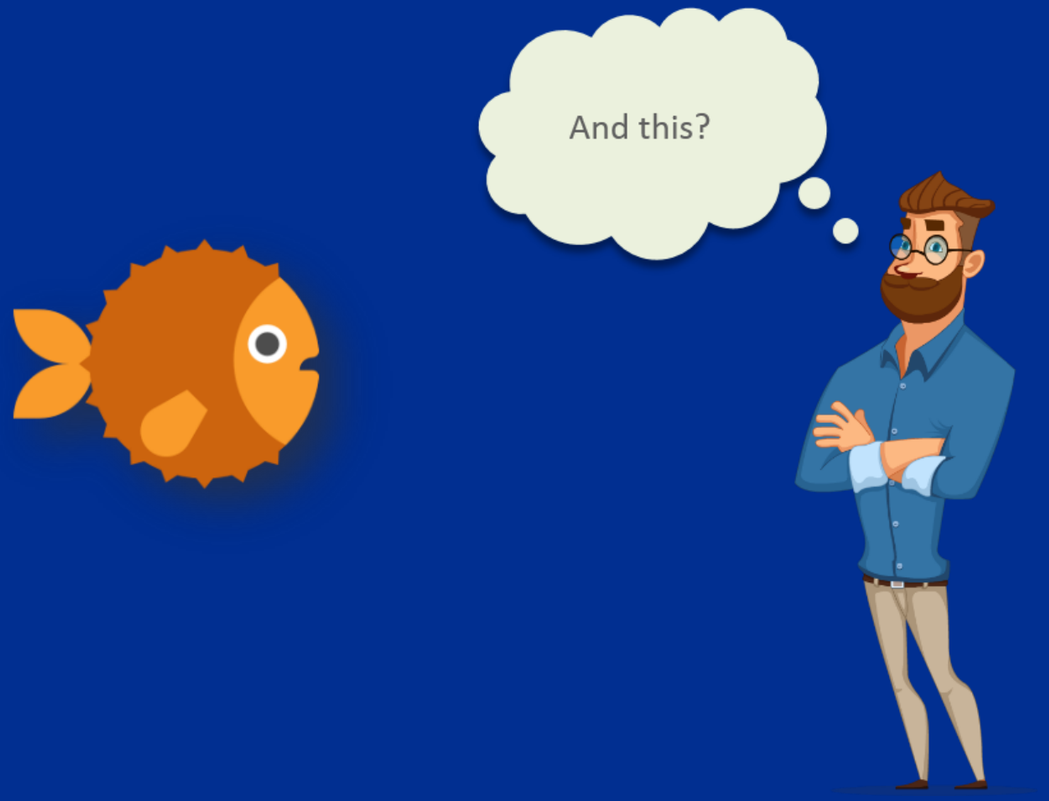
source: https://www.quora.com/What-are-prerequisites-to-start-learning-machine-learning

Machine Learning



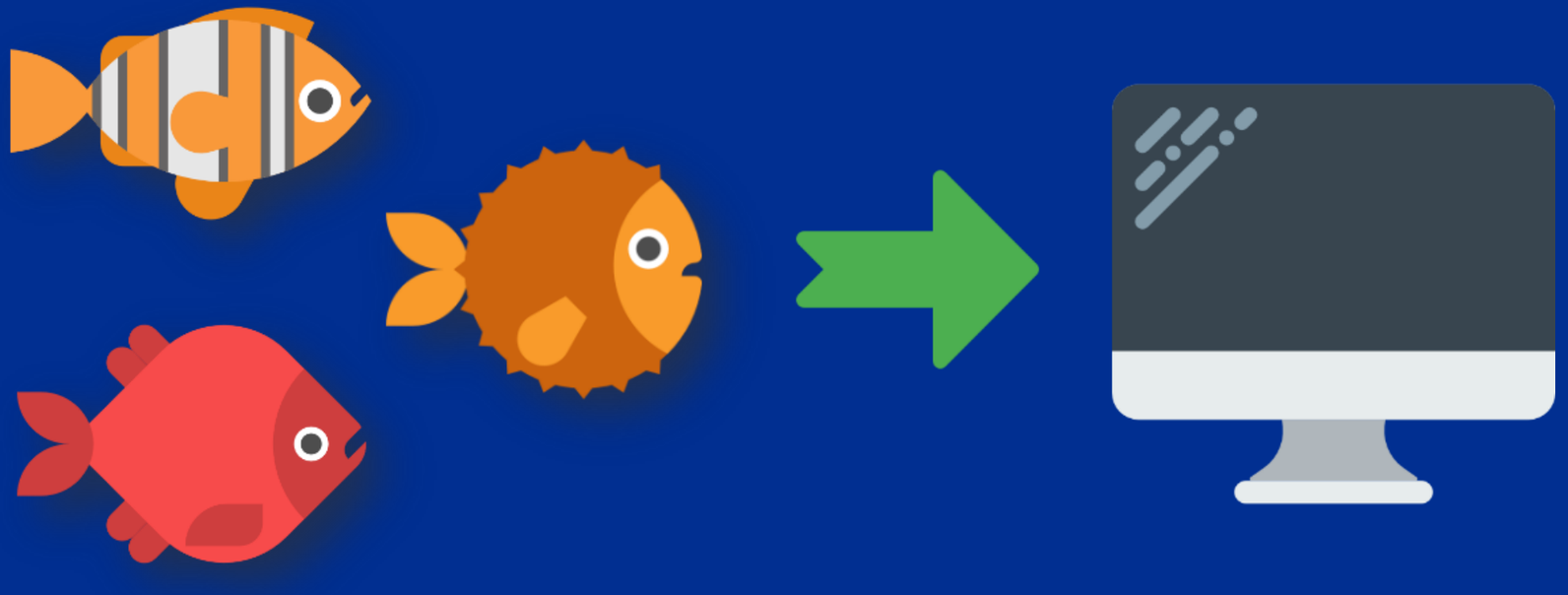
source: https://www.quora.com/What-are-prerequisites-to-start-learning-machine-learning

Machine Learning



source: https://www.quora.com/What-are-prerequisites-to-start-learning-machine-learning

Machine Learning



Input data

Machine

Source: https://www.quora.com/What-are-prerequisites-to-start-learning-machine-learning

Machine Learning

New Data

New Data

Machine

Machine

Result

Result

Source: https://www.quora.com/What-are-prerequisites-to-start-learning-machine-learning

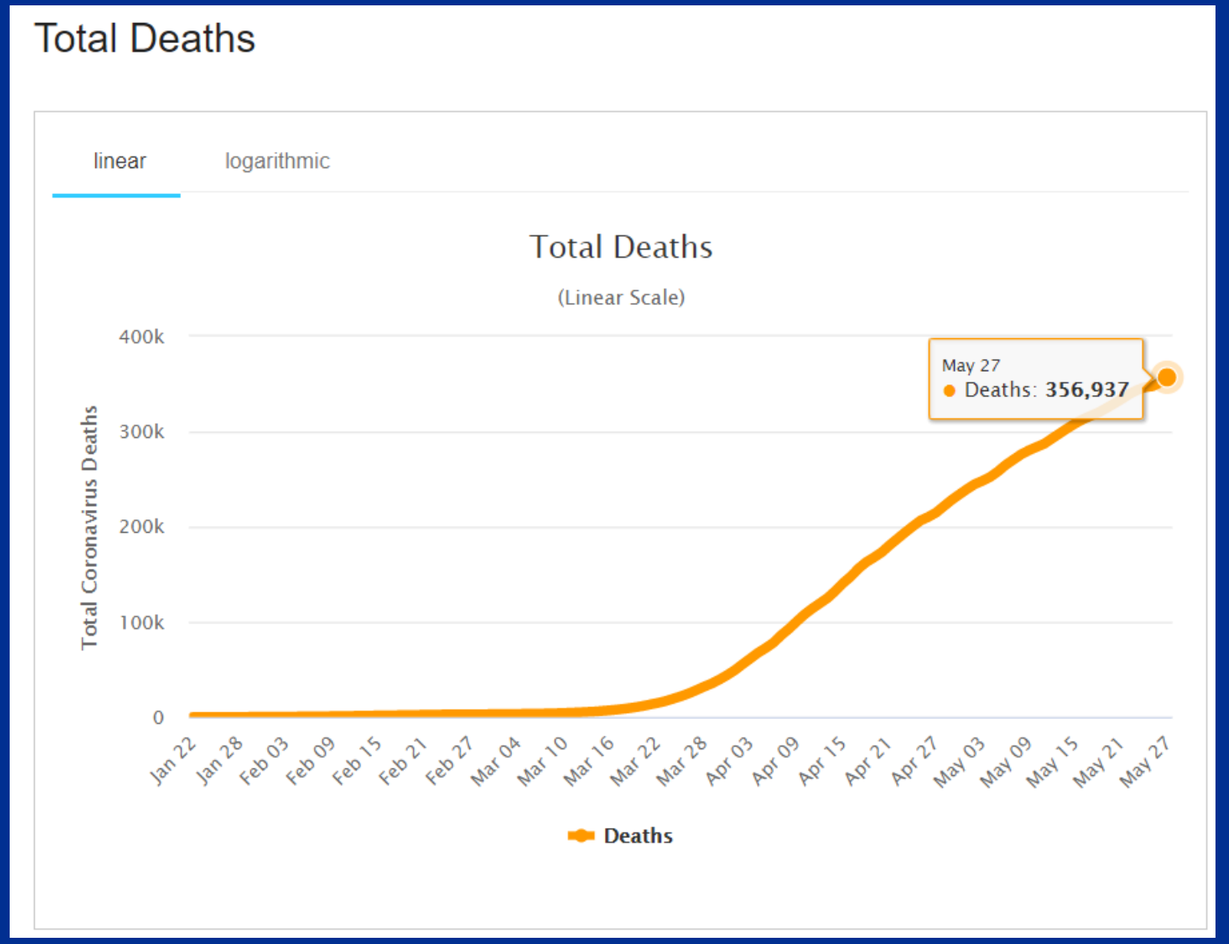
Pre-Requisites of ML

- A Pattern should exist
- Mathematical model/algorithm is unknown
- Abundant data

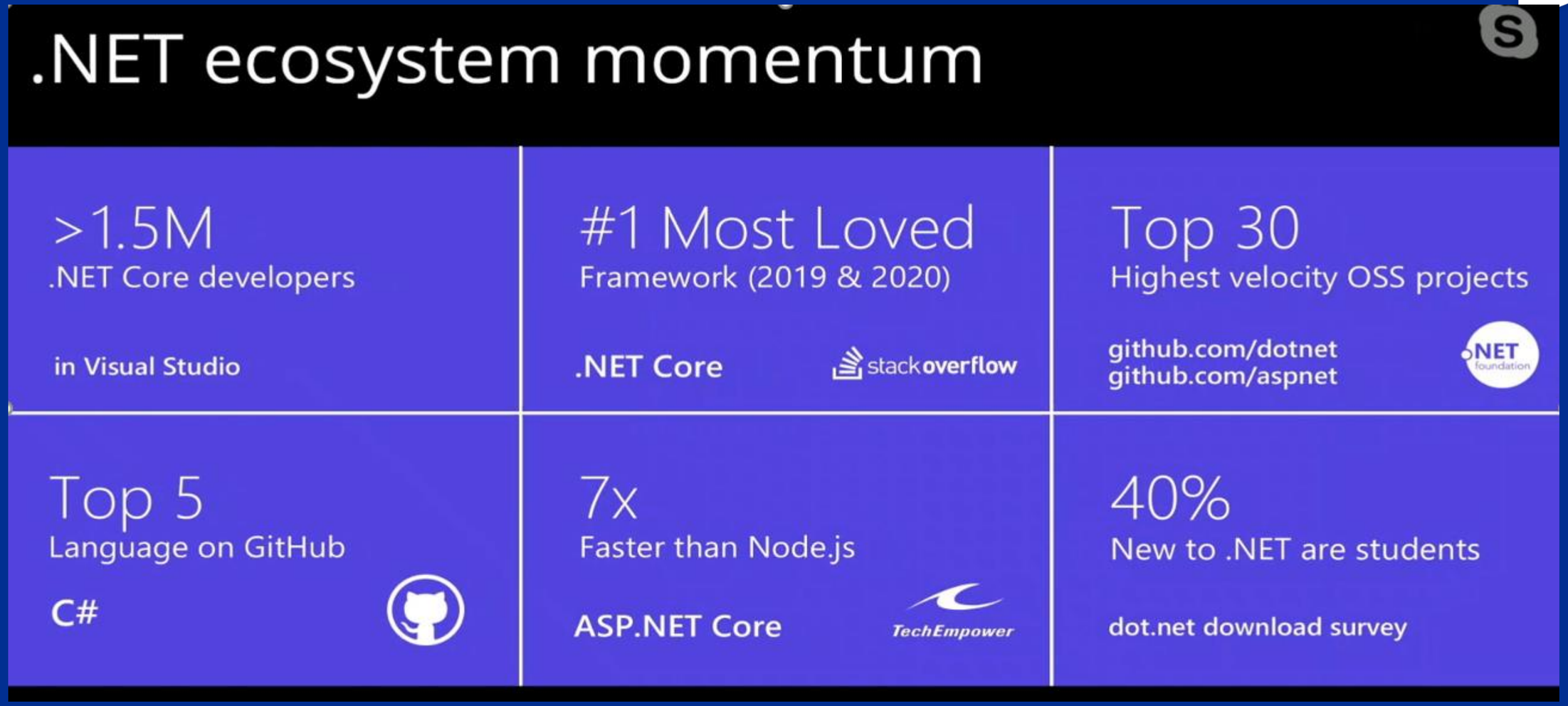
ML Frameworks

Predictions – Time Series Analysis

- The analysis of data organized across unit of time.
- It helps understand past trends
- Types: Trends, Seasonality, Irregularity, Cyclic



Why .NET ?

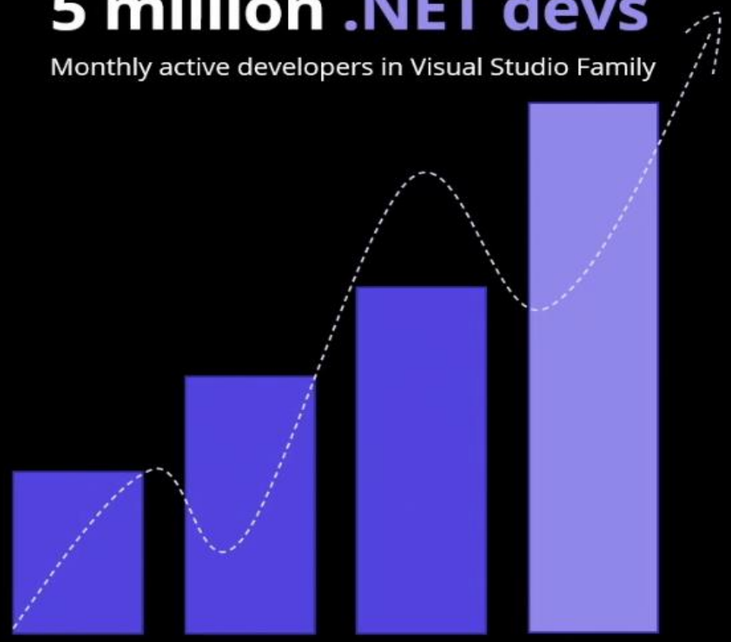

.NET ecosystem momentum
>1.5M
.NET Core developers
in Visual Studio
#1 Most Loved
Framework (2019 & 2020)
.NET Core
stackoverflow
Top 30
Highest velocity OSS projects
github.com/dotnet
github.com/aspnet
.NET foundation
Top 5
Language on GitHub
C#
7x
Faster than Node.js
ASP.NET Core
TechEmpower
40%
New to .NET are students
dot.net download survey

Why .NET ?

.NET adoption

5 million .NET devs

Monthly active developers in Visual Studio Family



Visual Studio — **+1 million new**

Monthly active .NET developers in the last year

.NET Core — **+600K new**

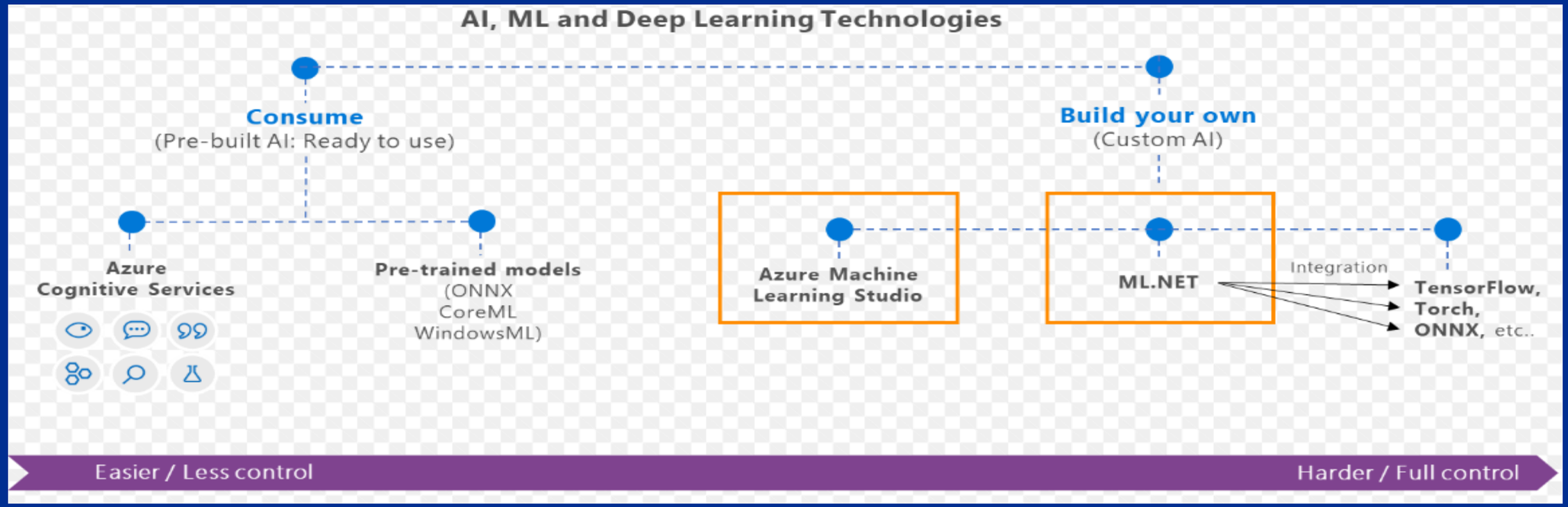
Monthly active .NET Core developers in the last year

.NET on Linux — **2 million**

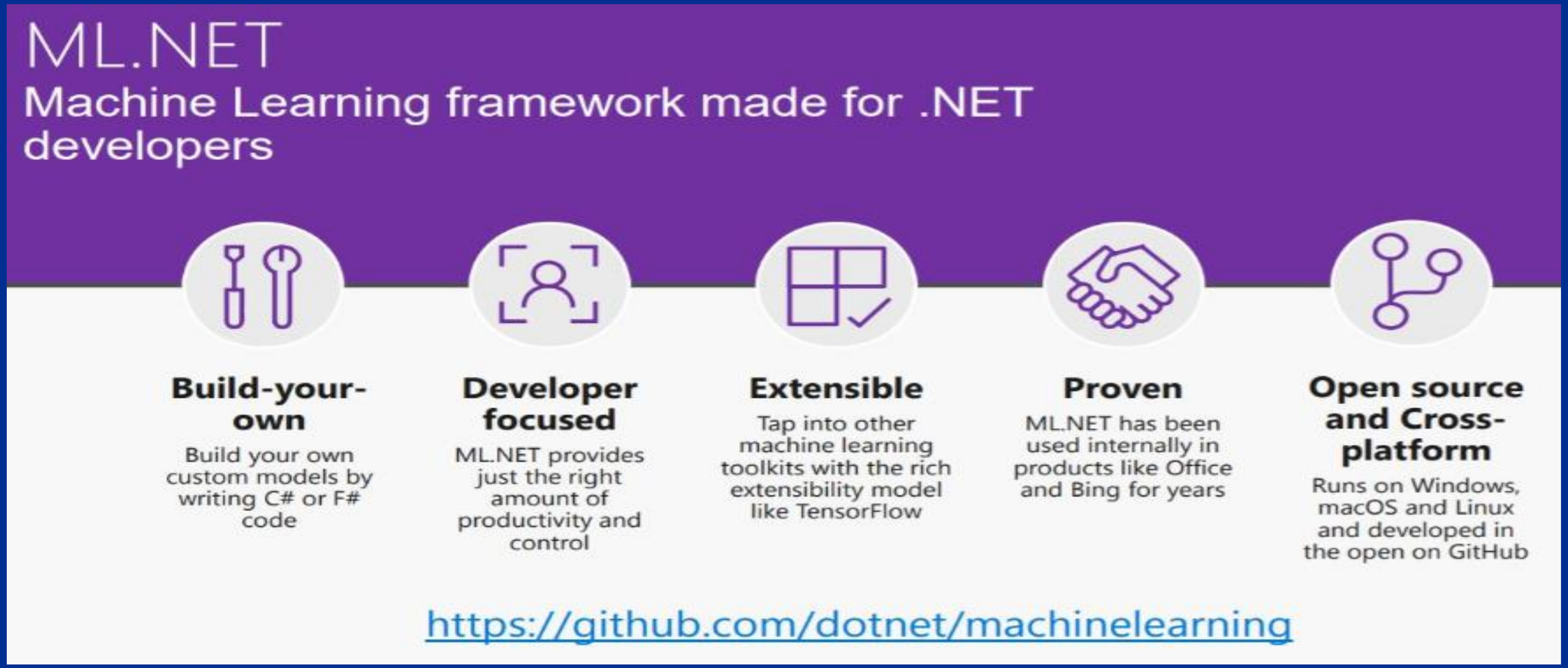
Publishes to Linux from Visual Studio

What is ML.Net ?

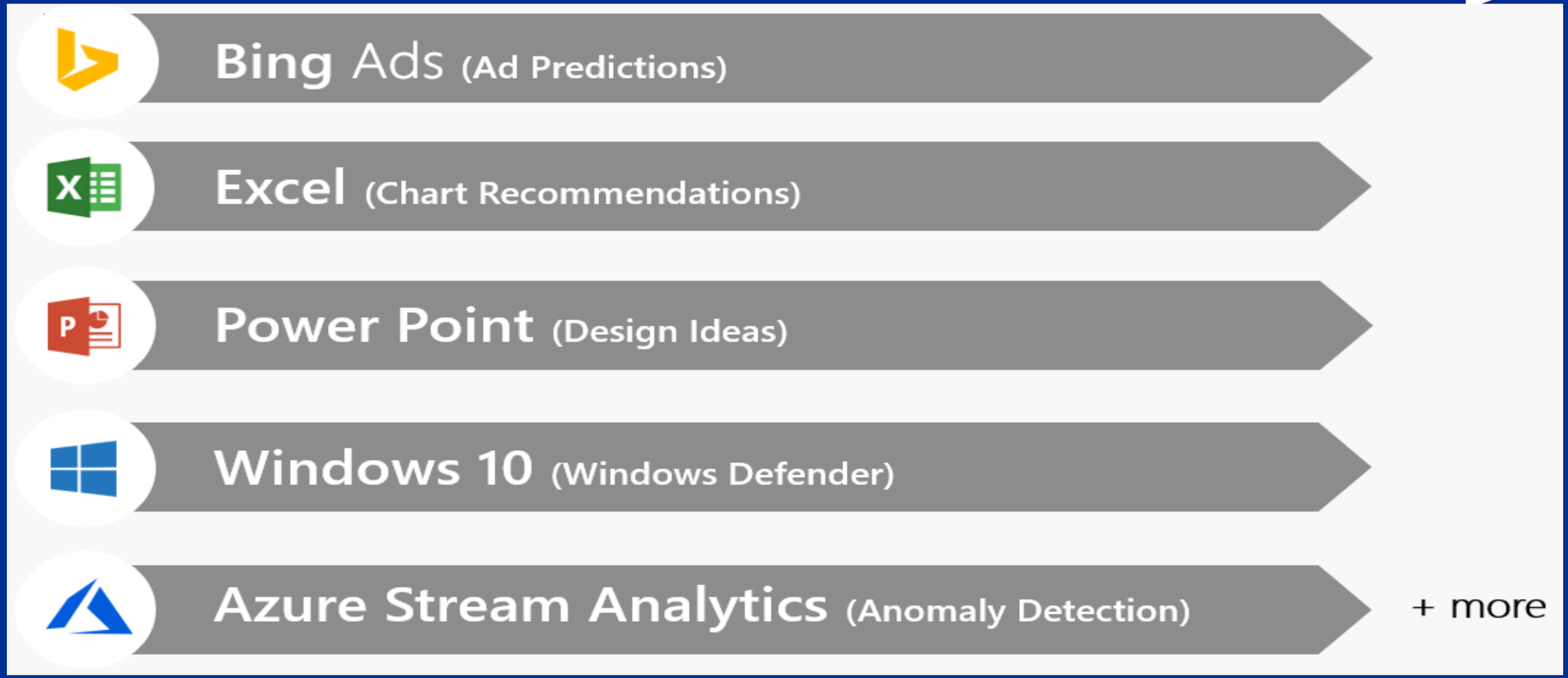
- ML Framework from Microsoft for developing custom AI/ML applications.
- Originated in 2002 as part of Microsoft Research Project



What is ML.Net ?



ML.Net – Proven at Scale, Enterprise ready



ML.Net – Possibilities

Sentiment Analysis

Forecasting

Issue Classification

Predictive maintenance

Image classification

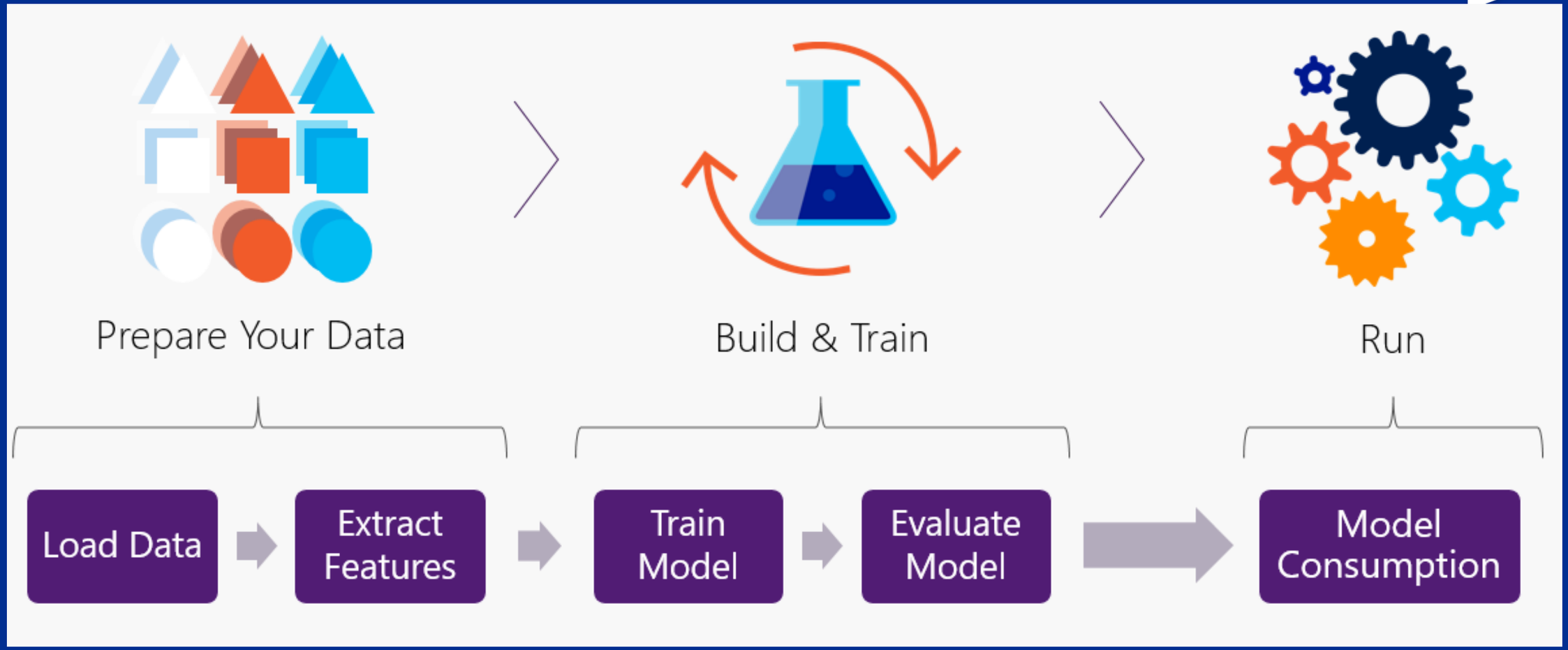
Recommendations

Object detection

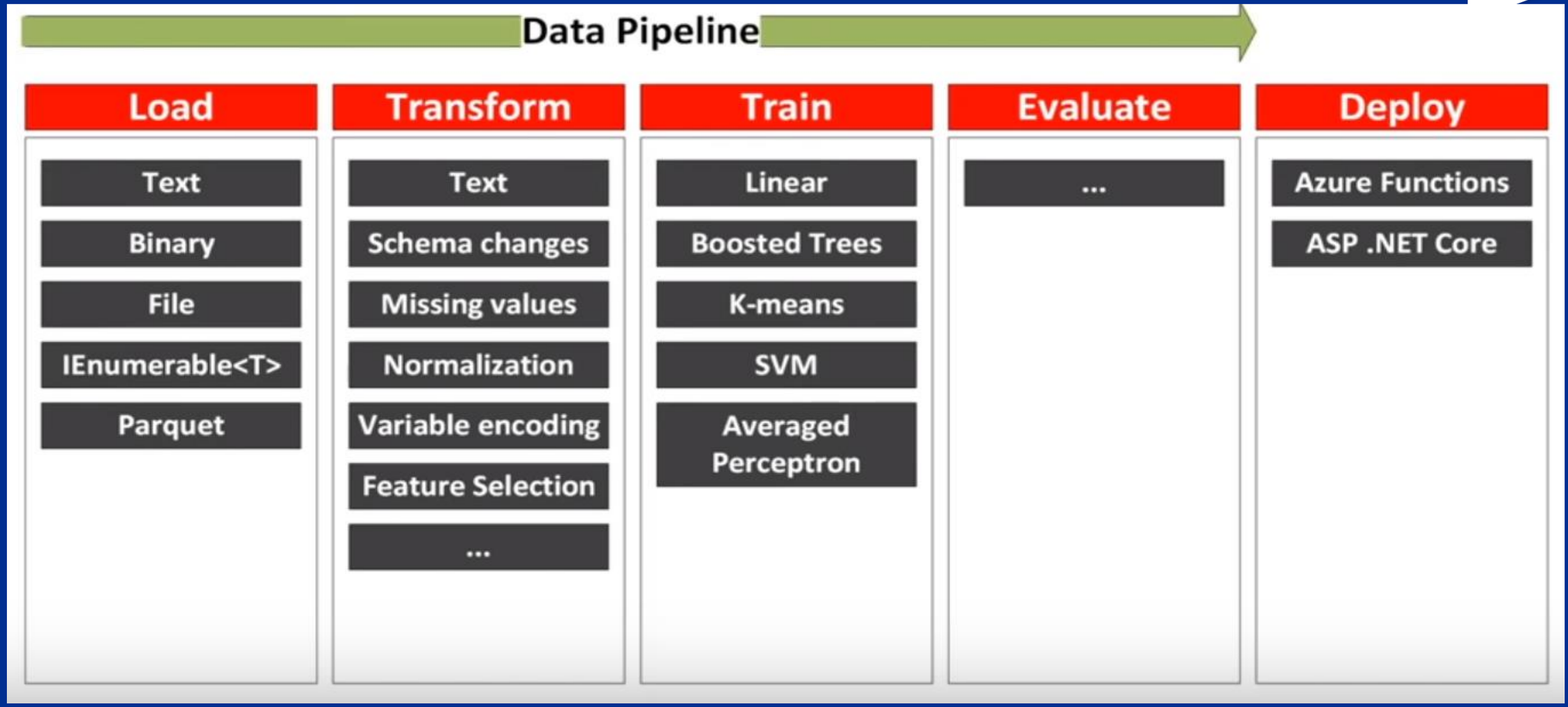
Customer segmentation

And more! Samples @ https://github.com/dotnet/machinelearning-samples

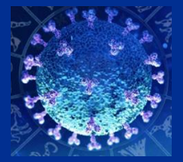
ML.Net – Workflow

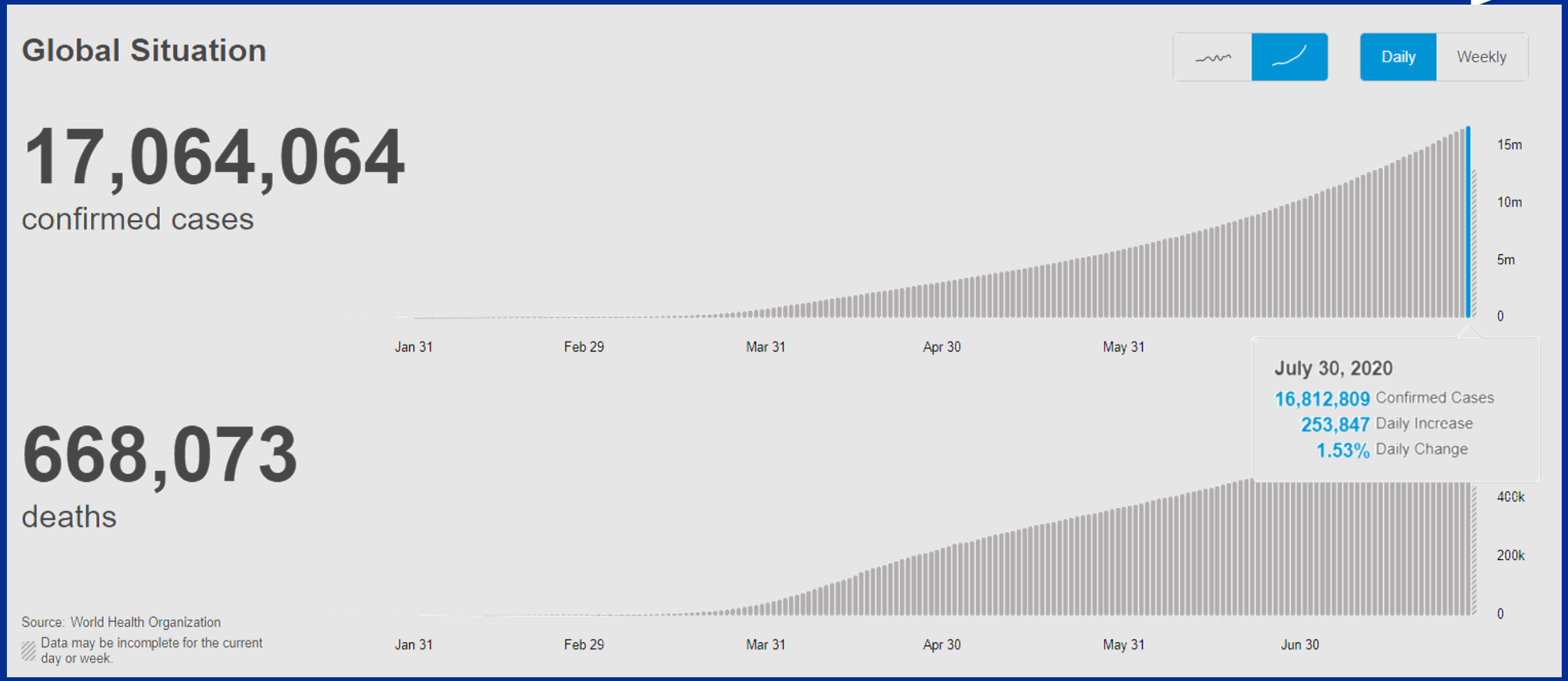


ML.Net – Data Pipeline



Covid-19 EDA and Prediction





#AzConfDev

Source: https://covid19.who.int/ Data as on July 30, 2020

EDA using .Net DataFrame

- A collection of columns
- In-memory representation of structured data
- Similar to Python Dataframe

- Load dataset from CSV
- Row/Column Selection
- Join/Merge DataFrame
- GroupBy
- Null Values

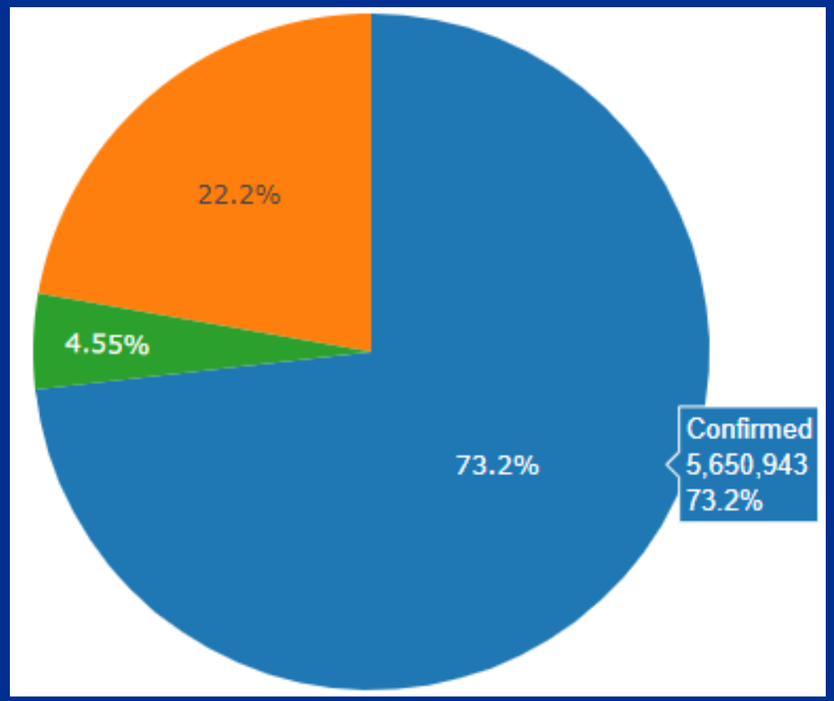
Machine Learning using ML.Net

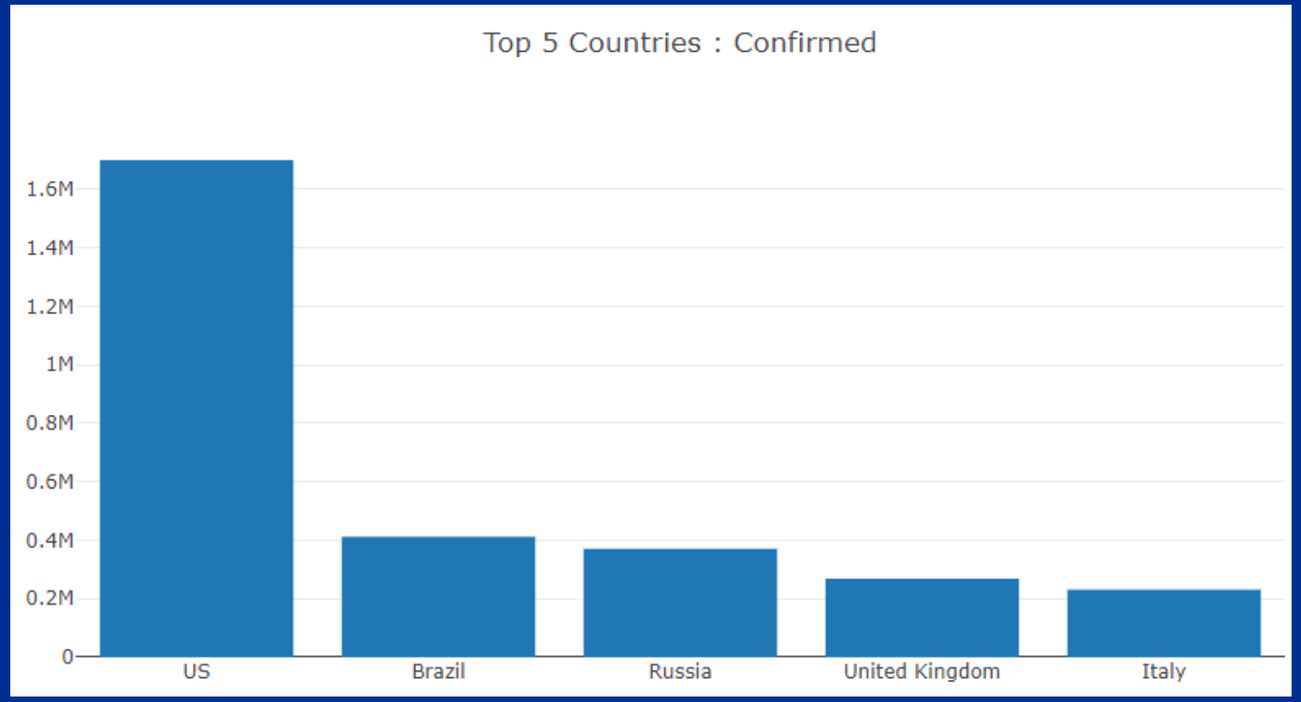
- Time Series
- Load Data – MLContext
- ML Pipeline
- Predict Confirmed cases in next 7 days.

Visualizations

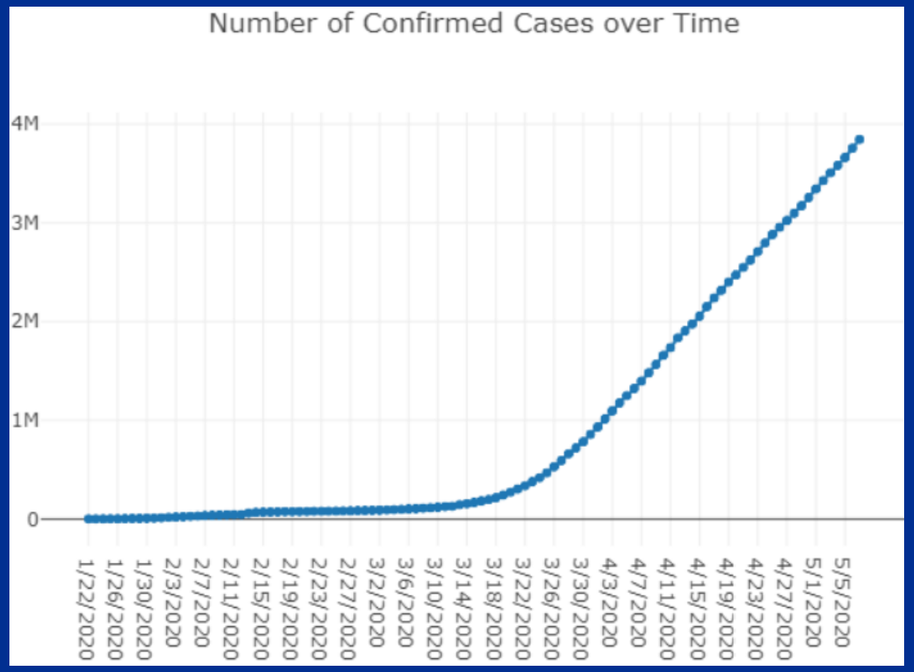
- Xplot – Plotly
- Bar
- Scatter
- Pie

Data Science – Data Frame

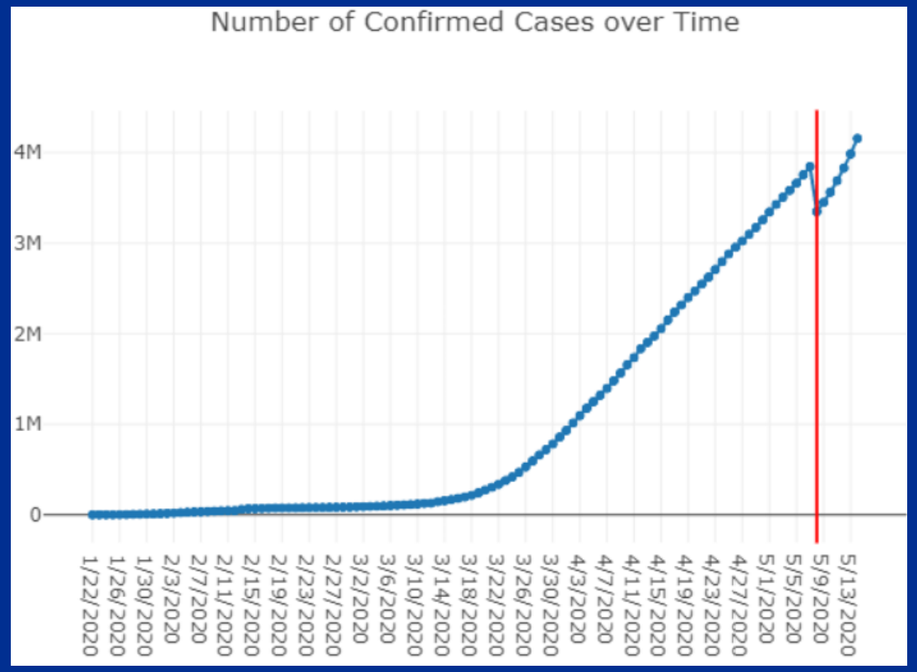




Machine Learning : ML.Net – Prediction (7 days)

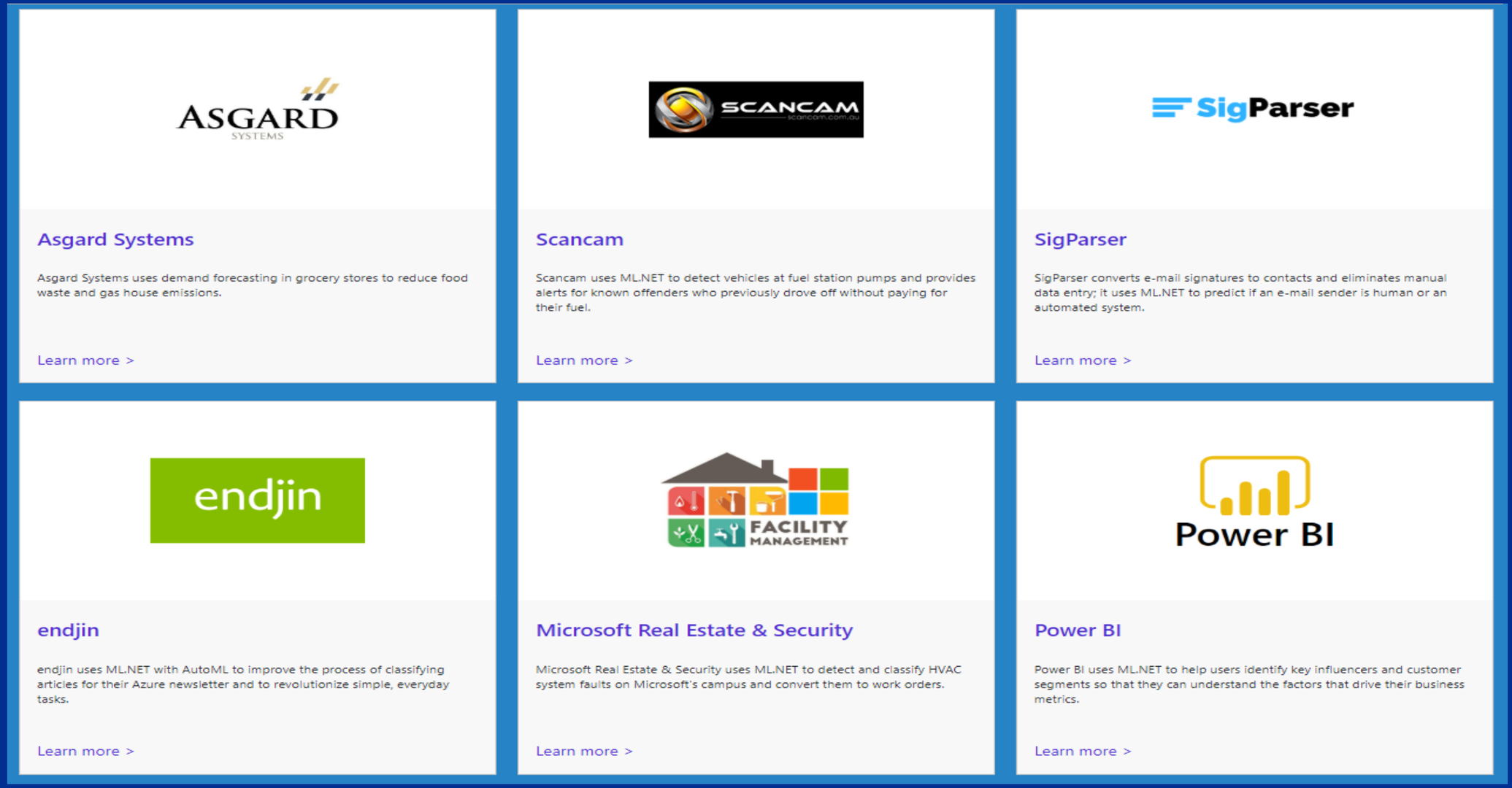


index	value
0	3348756
1	3450496
2	3563966
3	3690067
4	3830294
5	3985414
6	4156340



Demo

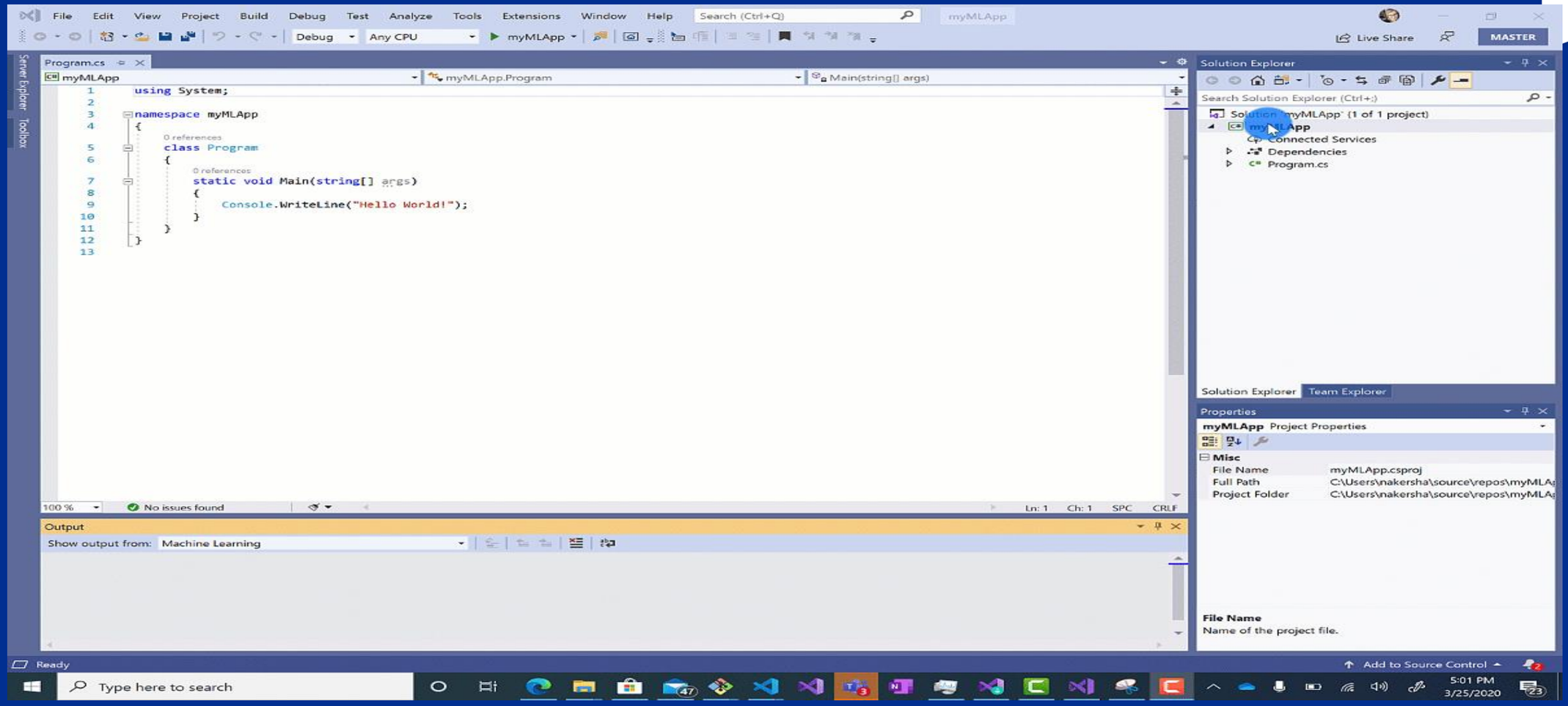
ML.Net : Customer Stories



#AzConfDev

https://dotnet.microsoft.com/apps/machinelearning-ai/ml-dotnet/customers

ML.Net : Model Builder(AutoML)



#AzConfDev

References

- ML.Net : https://dotnet.microsoft.com/apps/machinelearning-ai/ml-dotnet
- ML.Net: https://devblogs.microsoft.com/cesardelatorre/what-is-ml-net-1-0-machine-learning-for-net/
- ONNX : https://onnx.ai/
- Photo-Search (ONNX) : https://github.com/Tak-Au/Photo-Search
- Music Repair : https://www.youtube.com/watch?v=nnV-1q-z9uE
- ML Cookbook : https://github.com/dotnet/machinelearning/blob/master/docs/code/MlNetCookBook.md
- Deploy to Azure functions : http://luisquintanilla.me/2018/08/21/serverless-machine-learning-mlnet-azure-functions/
- https://rubikscode.net/2019/02/18/ultimate-guide-to-machine-learning-with-ml-net/
- https://www.youtube.com/watch?v=dojO4zEL9sg
- https://www.youtube.com/watch?v=zy7Y9CHji2k

https://dotnet.microsoft.com/apps/machinelearning-ai/ml-dotnet/customers

Resource

- Github: https://github.com/praveenraghuvanshi/tech-sessions/tree/master/21102022-AI-ML-using-ml-dotnet-azconf-2022

- Dev.to : https://dev.to/praveenraghuvanshi/covid-19-eda-and-prediction-using-net-dataframe-and-ml-net-c-introduction-nlb

https://dotnet.microsoft.com/apps/machinelearning-ai/ml-dotnet/customers

Platinum Partner

Gold Partner

Silver Partner

Q & A

Applying Data Science and Machine Learning using Microsoft .Net

Praveen Raghuvanshi

@praveenraghuvan

#AzConfDev