

Mixing Reality & AI: WinML, ML.NET e HoloLens



Gianni Rosa Gallina
R&D Specialist @ Deltatre
twitter: @giannirg
web: <http://gianni.rosagallina.com>

Milano – 15 Dicembre 2018

Questo lavoro è soggetto a Licenza **Creative Commons**



Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)



Tutti i nomi di prodotti o di aziende sono marchi registrati TM [®] dei rispettivi proprietari.
Il loro uso non implica alcuna affiliazione o sponsorizzazione.

Perché usare ML per risolvere un problema?

Alcuni problemi sono difficili da risolvere usando algoritmi o programmazione tradizionali.

Decidere se
un cliente
è affidabile o no

Prevedere
l'andamento e la
domanda dei clienti
in una stagione

Trovare un difetto
in un processo
industriale

Quando la soluzione è complessa da descrivere, descriviamo i dati.

Programmare l'improgrammabile

$f(x)$ { E' un cane?



$f(x)$ { Costo di una camicia?
"Ha bottoni stupendi...
con maniche lunghe... va
bene sia per uscire e sia
per lavoro"

Programmare l'improgrammabile



Tipico Processo di Machine Learning



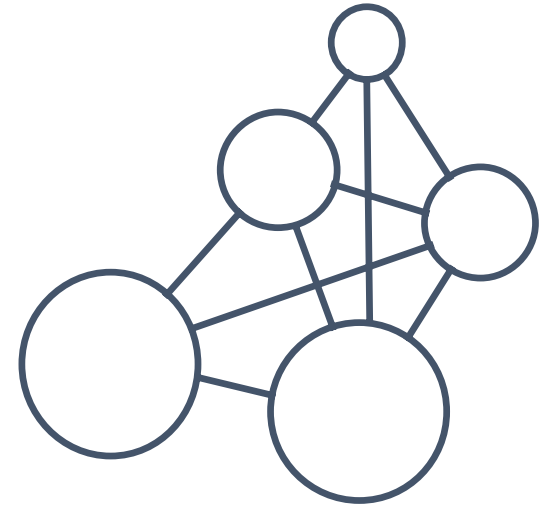
Microsoft AI Platform



Piattaforma AI per
sviluppatori Windows



Workflow completo
end-to-end



Windows come
prodotto AI-first

Windows ML

Dal Cloud all'Edge

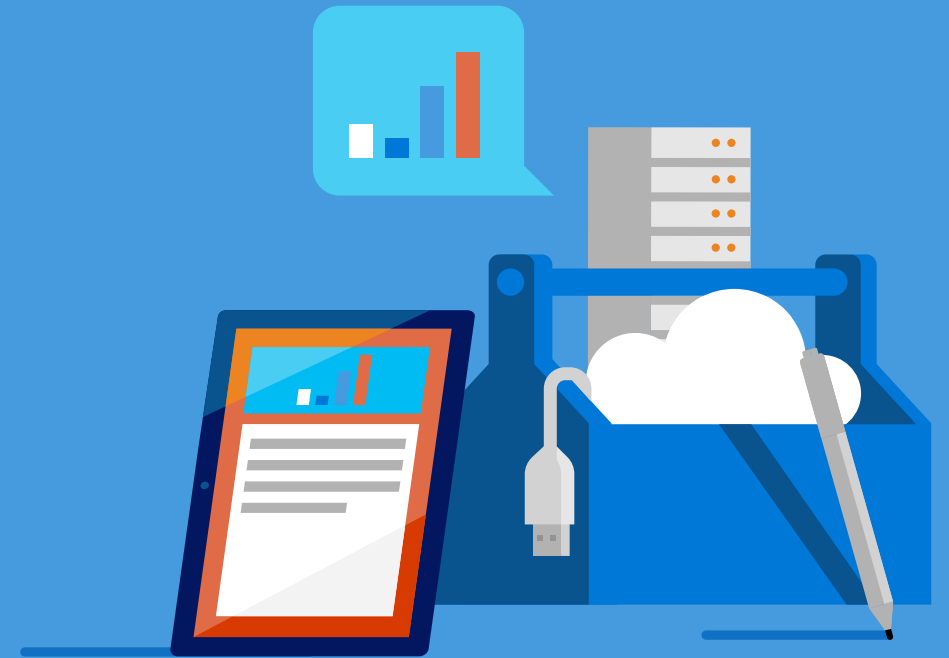
Crescita esplosiva

- disponibilità di dati aggregati
- potenza di calcolo (GPU, HW AI ad-hoc)

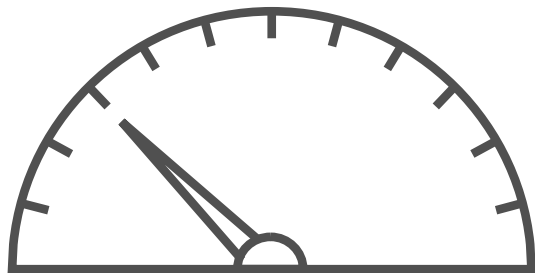
Efficienza nel Cloud → AI sui dispositivi

- Si *valutano* modelli creati in cloud
- Dimensioni ridotte dei dati a *runtime*
- Si sfruttano le GPU (inizialmente)
- Hardware ad-hoc per accelerare AI (in futuro)

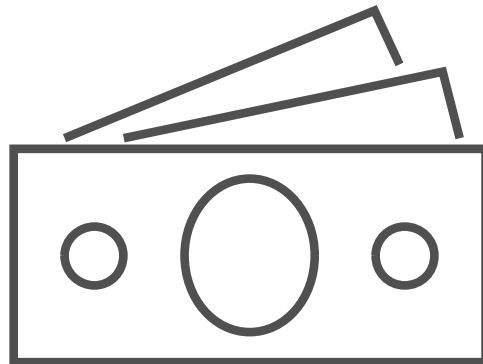
AI è pronta per il Cloud e l'Edge



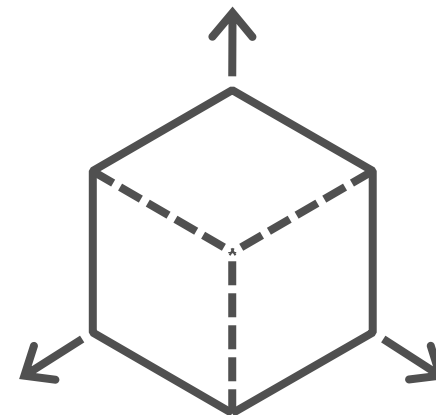
Al sui Dispositivi



Bassa latenza

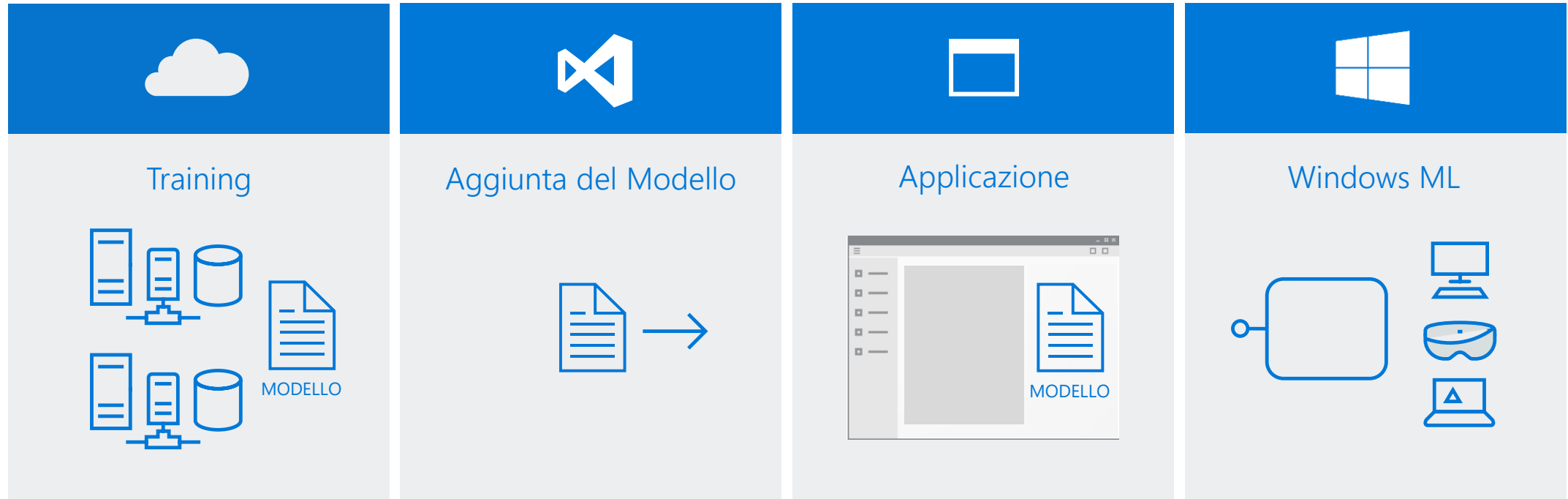


Scalabilità



Flessibilità

Dal Training all'Esecuzione



Training in Cloud

Azure Cognitive Services

- Servizi AI pre-definiti per *vision, speech, language, search* e molto altro
- Servizi *Custom Vision* permettono di personalizzare i modelli per casi d'uso specifici

Azure Machine Learning

- Ambiente *cloud-based* dove fare sviluppo, training, test, distribuzione, gestione e monitoring di modelli ML personalizzati
- I modelli si possono distribuire in cloud e/o su dispositivi Windows 10

Azure Batch AI

- Servizio per fare training dei modelli su cluster gestiti di GPU, per grandi basi dati che in genere richiedono molto tempo per il training.

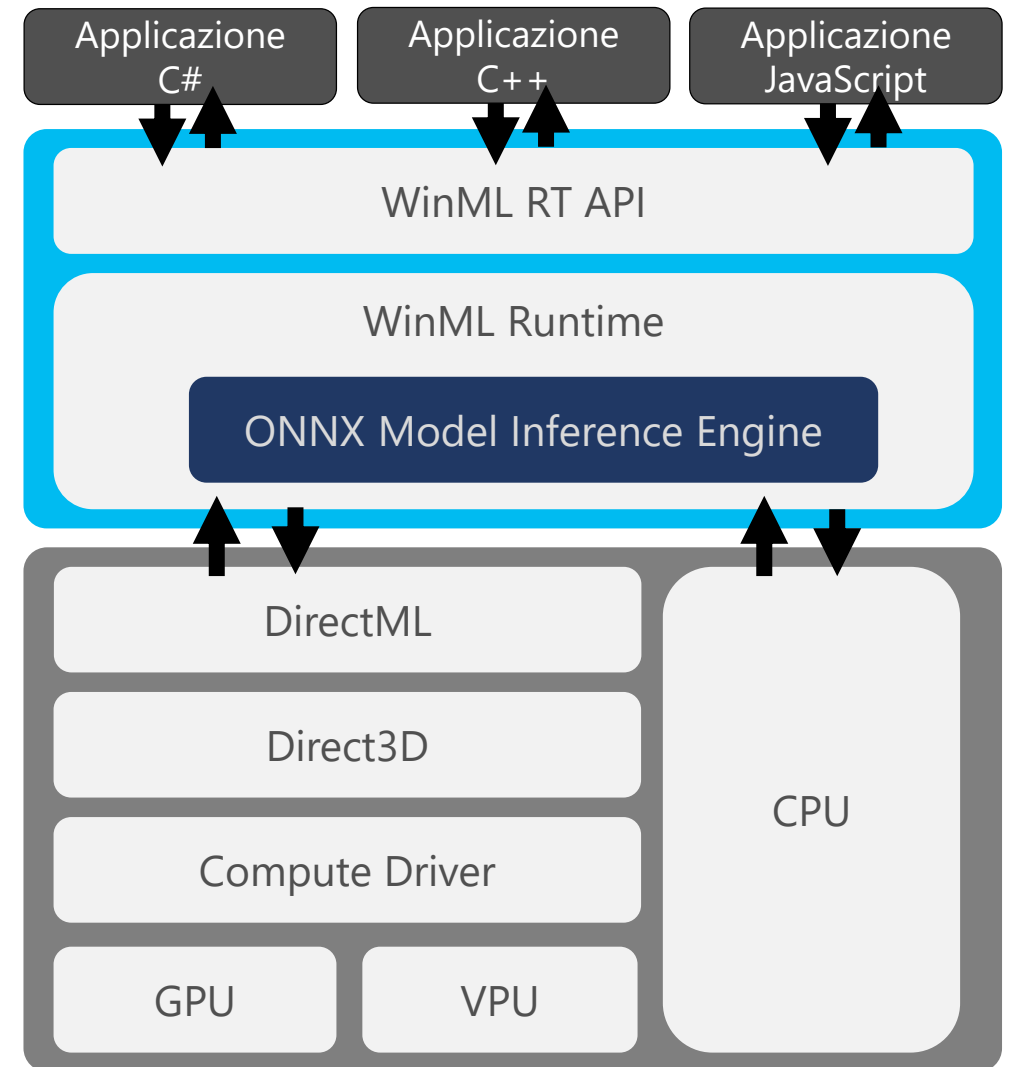
Esecuzione sui Dispositivi Windows

Accelerated Compute Layer

- Ottimizzato per le istruzioni CPU (fino a AVX-512)
- DirectML: implementazione per DX12
- MetaCommands permettono ottimizzazioni da parte dei produttori
- Supporto per FP16 riduce le dimensioni

Driver Model

- Supportato su tutti i dispositivi Windows 10
- WDDM-based per la gestione delle risorse
- Pronto per la Virtualizzazione
- Permette di scalare oltre le GPU



Piattaforme Windows ML

Windows 10 Desktop (October 2018)

Win32 e UWP

C++, C# e JavaScript (PWA)

Windows Server 2019

Richiede l'opzione "Desktop Experience"

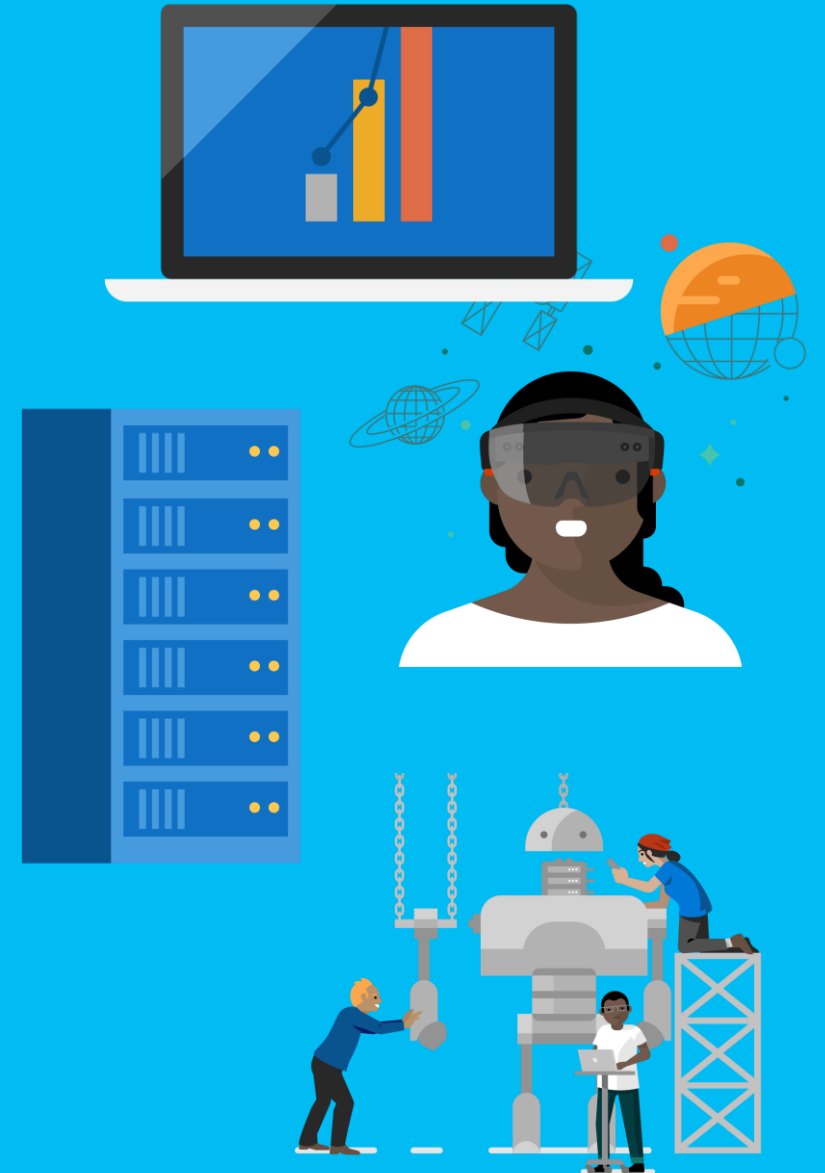
IoT

IoT Enterprise e IoT Core

Containers con accelerazione hardware

HoloLens (October 2018)

Supporto accelerazione hardware



Windows ML in Azione

Windows 10 Foto App

Tagging automatico per la ricerca

In media, ~88M di utenti attivi al giorno

Esecuzione sui dispositivi → *Privacy-friendly*

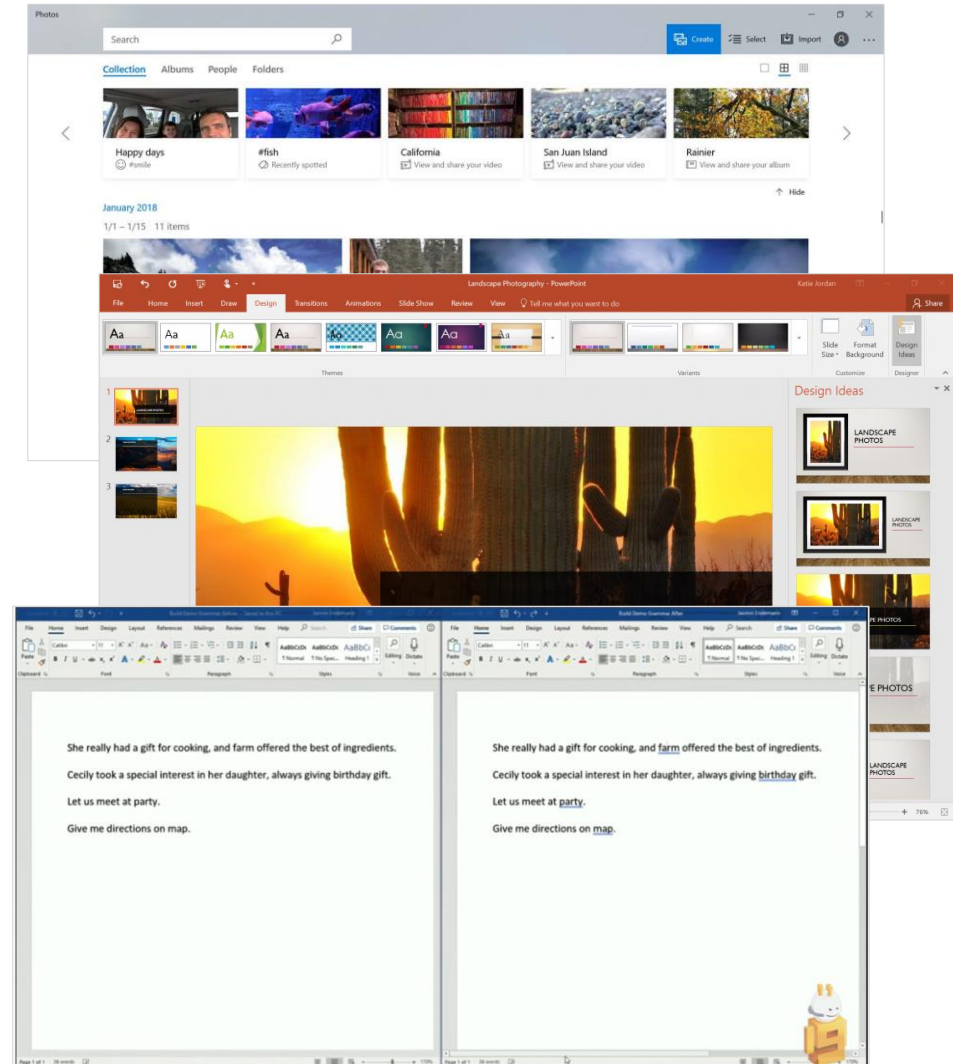
Microsoft PowerPoint

Esperimento per sostituire CNTK con WinML

Risparmio stimato di 70 VM con i carichi attuali
(4M slide/giorno per ~320K utenti)

Microsoft Word

Modello grammaticale / correttore ortografico



Lavorare con i Modelli ML

Windows usa modelli ONNX v1.2+

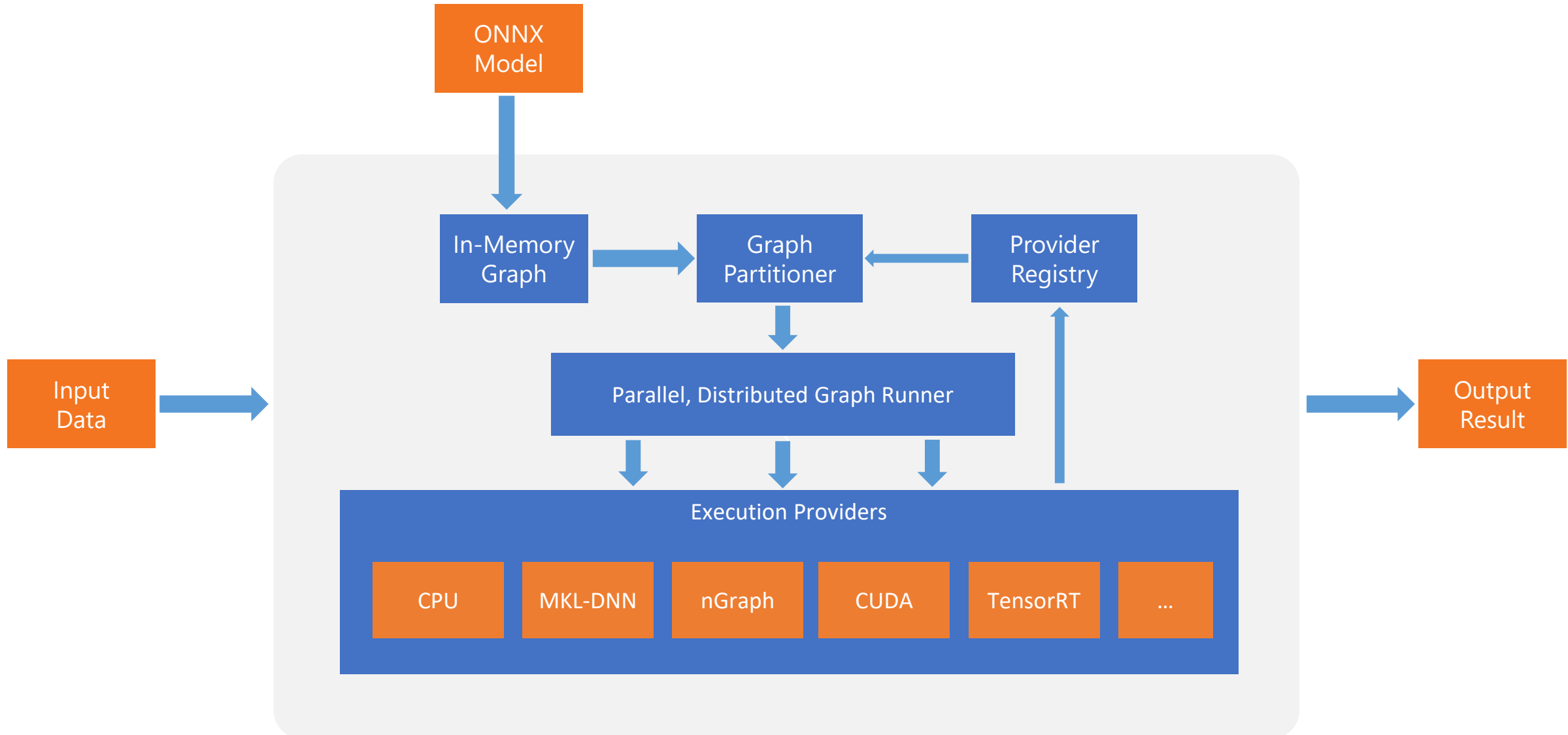
Open Neural Network Exchange, <https://github.com/onnx>

Formato standard, *open*, per definire i modelli ML (Microsoft, Facebook, Amazon e altri)



ONNX Runtime

<https://github.com/microsoft/onnxruntime>



Come ottenere modelli ONNX



ONNX Model Zoo

<https://github.com/onnx/models>



Azure Custom Vision

<https://customvision.ai>



Convertire modelli esistenti

WinMLTools: Caffe2, Pytorch, CoreML, SciKit, XGBoost, LibSVM, Keras, TensorFlow, CNTK, ecc.



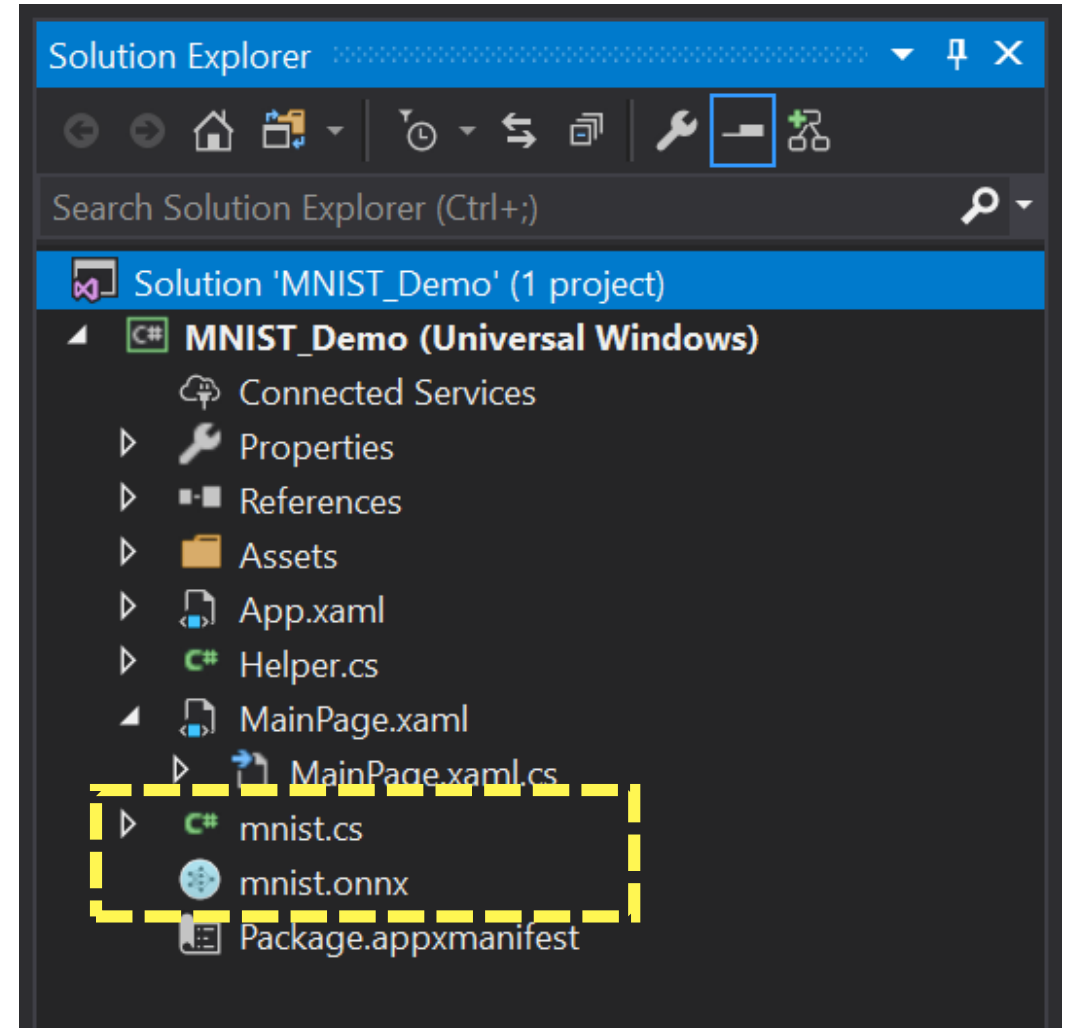
Modelli ad-hoc

Azure Machine Learning, ML.NET o altri framework ML/DL

Strumenti di Sviluppo

Visual Studio 2017 (v15.7+)

Modelli ONNX sono riconosciuti e gestiti con *MLGen*, per creare in automatico *wrapper* pronti all'uso



Strumenti di Sviluppo

Visual Studio 2017 (v15.7+)

Modelli ONNX sono riconosciuti e gestiti con *MLGen*, per creare in automatico *wrapper* pronti all'uso

Netron

Proprietà modelli, metadata Input/Output

Visualizzazione del grafo operatori

<https://github.com/lutzroeder/Netron>

MODEL PROPERTIES

format

ONNX v3

producer

CNTK 2.5.1

domain

ai.cntk

imports

ai.onnx v7

Graph

name

CNTKGraph

operators

Total: 12

Inputs

Input3

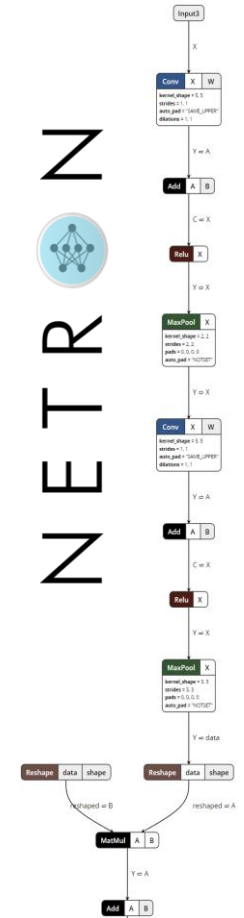
type: float32[1,1,28,28]

denotation: Image(Gray8,SRGB,NominalRange_0_255)

Outputs

Plus214_Output_0

type: float32[1,10]

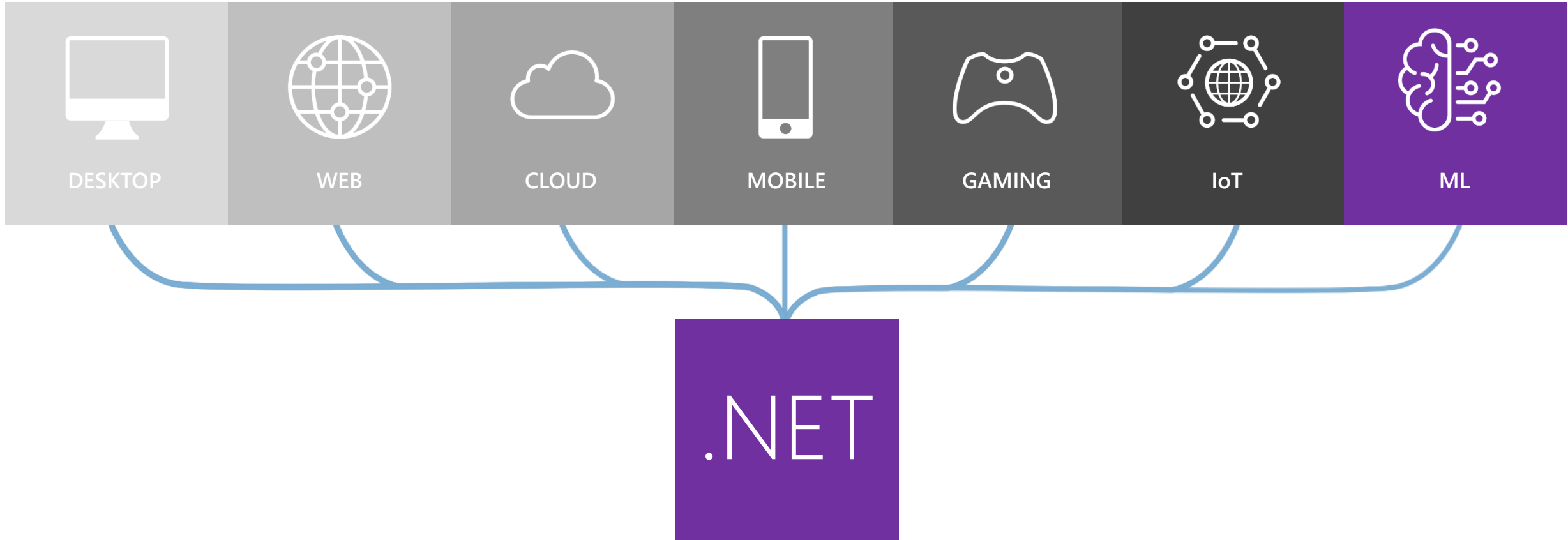


Demo

Modello ONNX in UWP App
(Object Detection – YOLOv3 Lite)

ML.NET

.NET: piattaforma per sviluppare qualunque cosa

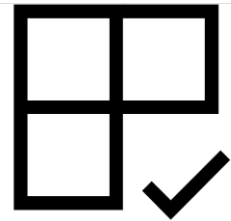


ML.NET: Framework per Modelli ML ad-hoc

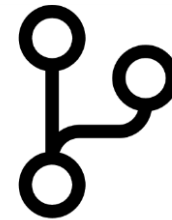
Supportato su Windows, Linux e macOS (.NET Core)



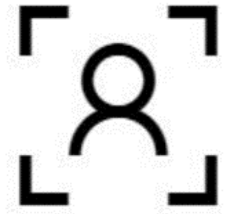
Sviluppo dei
propri modelli



Stabile ed
estendibile



Open-Source



Per sviluppatori .NET

ML.NET 0.8 è disponibile @ <https://github.com/dotnet/machinelearning>

API Machine Learning per sviluppatori

Training & Predizioni

Trasformazioni

Testo

Schema

Valori mancanti

Feature Categoriche

Normalizzazione

Selezione Feature



Algoritmi

Linear

Boosted Trees

Svm

K-Means

...



Strumenti

ML Data framework

Valutatori

Calibratori

Caricamento dati

...



Estensioni



ONNX



Scalabile e Pronto per Uso in Produzione



Bing Ads (Predizione annunci)



Excel (Suggerimento grafici)



Power Point (Suggerimento idee di layout)



Windows 10 (Windows Defender)



Azure Stream Analytics (Riconoscimento anomalie)

+ altri

Cosa si può fare



Sentiment Analysis



Previsioni



Classificazione elementi



Manutenzione predittiva



Classificatione immagini



Raccomandazioni



Riconoscimento oggetti

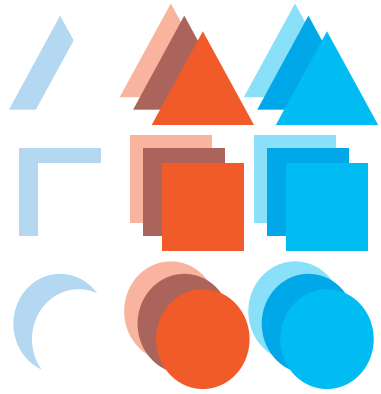


Segmentazione clienti

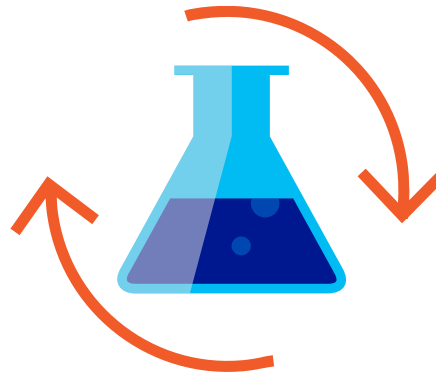


E molto altro! Esempi: <https://github.com/dotnet/machinelearning-samples>

Machine Learning Workflow



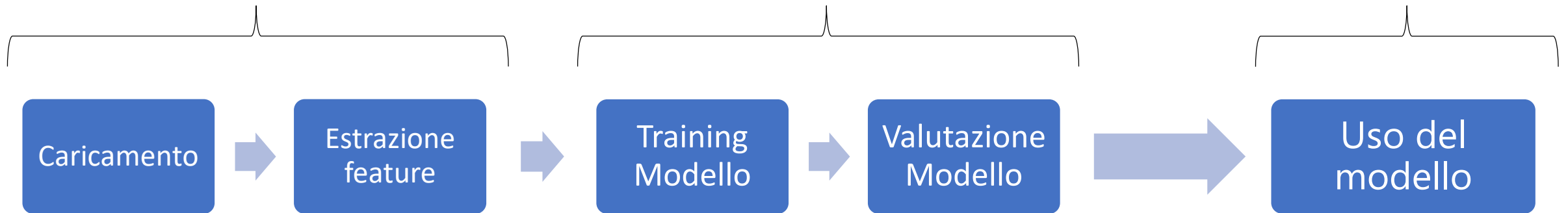
Preparazione Dati



Sviluppo & Training



Esecuzione



Demo

ML.NET Recommendation Engine

Custom Vision + ONNX +
WinML + HoloLens
=
Mixing Reality & AI

Demo

Custom Vision ONNX

HoloLens Object Classification

HoloLens Object Detection

Risorse e Riferimenti

Windows ML

- Per maggiori informazioni aka.ms/WindowsAIPlatform
- Esempi: <https://github.com/Microsoft/Windows-Machine-Learning>

Training in Azure

- Servizio Azure Custom Vision: customvision.ai
- Azure ML: <https://docs.microsoft.com/en-us/azure/machine-learning/service/>

ONNX

- Model Zoo: <https://github.com/onnx/models>
- Strumenti ONNX: onnx.ai/supported-tools
- Strumenti WinML: pypi.org/project/winmltools

ML.NET

- <http://dot.net/ml>
- Esempi: [http://aka.ms/mlnetsamples](https://aka.ms/mlnetsamples)



Risorse e Riferimenti

Demo di riferimento

- <https://elbruno.com/2018/06/29/winml-how-to-create-a-windows10-app-using-yolo-for-object-detection-2-of-n/>
- <https://blogs.msdn.microsoft.com/appconsult/2018/11/15/23535/>
- <https://blogs.msdn.microsoft.com/appconsult/2018/05/23/add-a-bit-of-machine-learning-to-your-windows-application-thanks-to-winml/>
- <https://blogs.msdn.microsoft.com/appconsult/2018/11/06/upgrade-your-winml-application-to-the-latest-bits/>
- https://github.com/dotnet/machinelearning-samples/tree/master/samples/csharp/end-to-end-apps/Recommendation-MovieRecommender/MovieRecommender_Model
- <http://dotnetbyexample.blogspot.com/2018/11/using-azure-custom-vision-object.html>
- <https://github.com/gianni-rg/global-ai-bootcamp-2018-demo>

Grazie!

Domande e Risposte

Contatti



Microsoft
Specialist

Programming in C#
Programming in HTML5
with JavaScript & CSS3

Microsoft
CERTIFIED

Solutions Developer
Windows Store Apps Using C#
Web Applications



Ing. Gianni ROSA GALLINA

R&D Specialist, Senior Software Engineer @ Deltatre

- Virtual/Augmented/Mixed Reality
- AI, Machine Learning
- Immersive video streaming e grafica 3D per eventi sportivi
- Cloud solutions, web backends, serverless, video workflows
- Mobile apps dev (Windows / Android / Xamarin)
- End-to-end solutions con Microsoft Azure

gianni@rosagallina.com



@giannirg



PLURALSIGHT

Author

<http://gianni.rosagallina.com>