

Machine learning za .NET developere



Connecting the dots

Neum 27-29/3/2019

Adnan Dželihodžić, PhD
Vodeći IT projektant@CBBH

NetWork 

NetWork

powered by:



platinum sponsor

LANACO
INFORMACIONE TEHNOLOGIJE

platinum sponsor

LOGOSOFT

telekomunikacijski sponsor

HT ERONET

oficijelni brend konferencije

Lenovo

gold sponzori

Addiko Bank



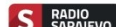
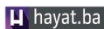
silver sponzori



prijatelji konferencije



medijski sponzori



Agenda

- Uvod u ML
- O ML.NET
- ML.NET glavne funkcionalnosti
- Demo
 - Klasifikacija
 - Regresija
 - Klasifikacija slika - TensorFlow



Machine learning



„Machine learning is the science of getting computers to act without being explicitly programmed.” – Stanford

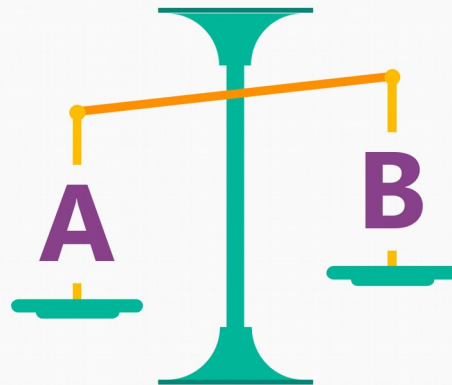
- 5 pitanja - > 5 kategorija problema → 5 vrsta algoritama
 - 1) Da li je ovo A ili B?
 - 2) Da li je ovo neuobičajeno (čudno)?
 - 3) Koliko mnogo ili koliki broj?
 - 4) Kako je ovo organizovano?
 - 5) Šta trebam uraditi sljedeće?



Klasifikacija

Is this A or B?

Classification algorithms



Anomalije



Is this weird?

Anomaly detection algorithms





Regresija

How much? How many?

Regression algorithms

Monday



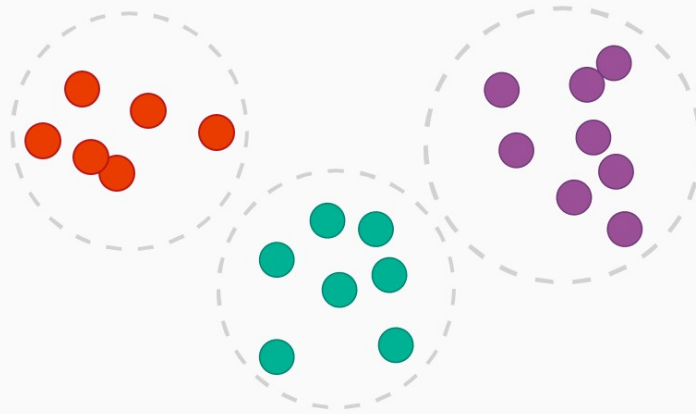
72°

Tuesday



How is this organized?

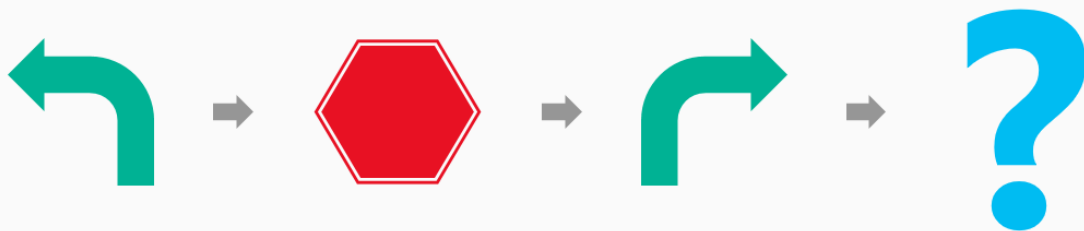
Clustering Algorithms





What should I do now?

Reinforcement Learning Algorithms



ML tok



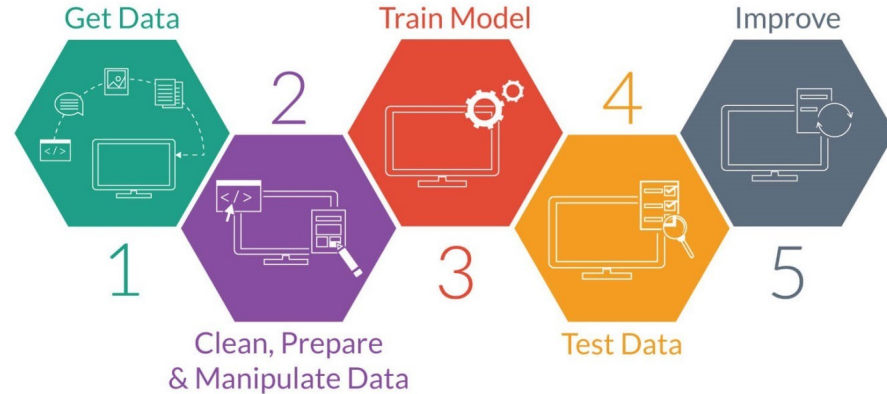
Prepare Your Data



Build & Train



Run



ML.NET





Šta je ML.NET?

- „*ML.NET is a cross-platform open-source machine learning framework which makes machine learning accessible to .NET developers.*”
- Cross-platform
 - Windows
 - Linux
 - macOS
- Mogućnost pravljenja modela, proširivost, Open Source
- **Upotreba u postojećim .NET aplikacijama**

ML.Net historija



- Originalno razvijeno od strane Microsoft Research odjela za internu upotrebu
- Primjeri: Excel chart recommendations, Bing features, Windows defender
- Inicijalna verzija objavljena u maju 2018. (0.1)
- Trenutna verzija 0.11 objavljena 5.3.



Najznačajnije funkcionalnosti

- Učitavanje i transformacija podataka
- Supervizijsko učenje
- Nesupervizijsko učenje
- Evaluacija modela
- Objašnjenje modela
- Korištenje prethodno treniranih modela (ONNX, TensorFlow)
- Pipeline učenjac



Najznačajnije funkcionalnosti

Developer friendly APIs for Machine Learning

Training

Consumption



Extensions



Transforms

Natural Text

Schema

Missing values

Categorical

Normalization

Feature Selection



Learners

Linear

Boosted Trees

Svm

K-Means



Misc.

ML Data framework

Evaluators

Calibrators

Data loaders





Instalacija

- Potreban .NET SDK!

Install-Package Microsoft.ML -Version 0.11.0

```
dotnet add package Microsoft.ML --version 0.11.0
```

Učitavanje podataka



- ML.NET može učitati sljedeće formate podataka:
 - Tekst (CSV)
 - Parquet
 - Binary
 - `IEnumerable<T>`
 - File sets



Transformacija

- Transformacije teksta
- Izmjena šeme podataka
- Upravljanje nedostajućim vrijednostima
- Kodiranje kategorijskih varijabli
- Normalizacija
- Izbor relevantnih atributa za trening (najznačajnijih)

- Linear (e.g. SymSGD, SDCA)
- Boosted Trees (e.g. FastTree, LightGBM)
- K-Means
- SVM
- Averaged Perceptron



Evaluacija i deployment

- Omogućene metrike i evaluacije modela
- Mogućnost snimanja modela kao binarni file (npr. Zip)
- Mogućnost integracije modela u bilo koju .NET aplikaciju

Podržani slučajevi upotrebe



- Standardna klasifikacija / regresija
- Procesiranje teksta
- Procesiranje slika
- Detekcija anomalija
- Rangiranje
- Sistemi za preporučivanje
- Segmentacija / klastering



Šta još nedostaje?

- Deep learning (treniranje)
- Distribuirano treniranje
- Podrška za ARM/x86...

Demo 01 – Binarna klasifikacija

1. Učitavanje podataka
2. Transformacija podataka (grupisanje, enkoding, popunjavanje nedostajućih vrijednosti)
3. Dodavanje algoritma za učenje
4. Treniranje
5. Evaluacija modela
6. Snimanje modela na disk
7. Poziv modela
8. Testiranje modela



Demo 02-klastering

1. Učitavanje podataka i razdvajanje
2. Procesiranje podataka
3. Kreiranje i treniranje modela
4. Evaluacija algoritma
5. Snimanje modela
6. Testiranje na primjeru



Demo 03 – Klasifikacija slika - TensorFlow

1. Učitavanje pretreniranog modela (Inception)
2. Treniranje modela sa novim slikama
3. Testiranje modela



Linkovi

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

<https://docs.microsoft.com/en-us/dotnet/machine-learning/>

<https://github.com/dotnet/machinelearning-samples>

Moj GitHub repo:

<https://github.com/adnandz/network2019>





DA LI VAM SE SVIDJELO OVO PREDAVANJE?

Ocijenite ga i osvojite vrijedne nagrade!

Ocjenjivanje predavanja kojima ste prisustvovali je moguće na dva načina:

1. Posjetom web stranici

<https://www.networkkonferencija.ba/profile/evaluation> vršite evaluaciju u sklopu svog korisničkog profila.

2. Ocijenite predavanja kroz konferencijske mobilne aplikacije pod nazivom NetWork Conference koje su dostupne za Android i iOS mobilne uređaje.

U oba slučaja potrebno je prijaviti se sa korisničkim računom sa kojim ste se registrovali za učešće na konferenciji.

Ispunjavanjem upitnika, pomažete Organizatoru u organizaciji narednih konferencija.

