



Giovedì 29 Febbraio 2024

in presenza e in streaming

alle ore 19:00

#Kaggle #Competitions

@PyDataVenice #14 #Meetup #PyData



Fabio Dal Forno
RPA Engineer
@ Generali Italia



Alessandra Bilardi
Data / Automation Specialist
@ Corley Cloud



Giovedì 29 Febbraio 2024

in presenza e in streaming

alle ore 19:00

#Kaggle #Competitions

@PyDataVenice #14 #Meetup #PyData



Fabio Dal Forno
RPA Engineer
@ Generali Italia



Francesco Patanè
Research Fellow
@ Università di Padova



Giovedì 29 Febbraio 2024

in presenza e in streaming

alle ore 19:00

Motion Analytica

#Kaggle #Competitions

@PyDataVenice #14 #Meetup #PyData



Fabio Dal Forno
RPA Engineer
@ Generali Italia



Francesco Patanè
Research Fellow
@ Università di Padova

Promotori di PyData Venice #14



Motion Analytica

Insights from people & things in motion



Agenda

Speech

Prossimo incontro

Spritz





Kaggle Overview

Alessandra Bilardi - Data & Automation Specialist @ Corley Cloud

@PyDataVenice #14 #Meetup #PyData

Agenda

Kaggle in numbers
The platform

Kaggle in numbers

Kaggle history

Event	When	Notes
Foundation	April 2010	a Kaggle user was the President and Chief scientist (November 2010)
Acquiring	March 2017	Kaggle is under Google LLC
Over 1 million registered users	June 2017	
Introduced Models	February 2023	Users can use pre-trained models
Over 15 millions users	October 2023	Users are in 194 countries
Over 17 millions users	February 2024	







—

may






open source

with you

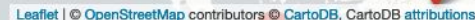
Kaggle sections

Section	Symbol & alternative name	Quantity
Competitions	 Competitions	26000
Datasets	 Datasets	298K
Models	 Models	2300
Code	 Notebooks	1M
Discussions	 Solution write-ups	4000
Learn	 Courses	70+ hours

Kaggle users

User	Symbol	Competitions sub / Users
Novices		119K / 16.85M
Contributors, notebook & upvote		69K / 232K
Experts, bronze medals		10K / 16K
Masters, gold medal or silver medals		2K / 2.6K
Grandmasters, gold medals or silver medals		322 / 497

Notebook Input Output Logs Comments (2)



<https://www.kaggle.com/code/steubk/kaggle-grand-masters-map>

Kaggle resources



Resource	Quantity
API (Competitions, Datasets, Notebooks)	3 sections
NVIDIA TESLA P100 GPUs	30h / week
Tensor Processing Units (TPUs)	20h / week
/kaggle/working/	20 GB
Organizations (Hub)	767

Kaggle is useless for ..



Useless	Vote
Entrare nel merito di architetture, MLOps e tutto ciò che trasforma un algoritmo standalone in qualcosa production-ready	2
Imparare a scrivere codice pulito e ben organizzato	5

Kaggle is useful for ..

Useful	Vote
Comprendere i dati e il problema a livello di business, oltre il solo ML	7
Imparare il tuning degli algoritmi	8
Crearsi un portfolio di veri progetti di AI/ML	8
Conoscere lo stato dell'arte degli algoritmi di AI/ML che davvero funzionano	9
Imparare il feature engineering, con accorgimenti e trucchetti vari che non trovate sui manuali universitari e sui libroni di statistica	10
Capire i concetti di overfitting, di validazione e generalizzazione in scenari complessi	10 e lode

The platform



The platform

[my work space](#)

- Competitions
 - [Forecasting](#)
 - [AI assistants](#)
- Grandmasters
 - [Alberto Danese](#)
 - [Kalilur Rahman](#)
- Datasets
 - [Meta Kaggle](#)
 - [COVID-QA](#)
- Models
 - [Mixtral](#)
 - [MIRNet](#)



Questions ?

@PyDataVenice #14 #Meetup #PyData



Kaggle Competitions

Fabio Dal Forno - RPA Engineer @Generali Italia

@PyDataVenice #14 #Meetup #PyData



Questions ?

@PyDataVenice #14 #Meetup #PyData

Prossimo incontro



- venerdì 19 aprile ore 19:00

Proposte





Thanks for listening.

@PyDataVenice #14 #Meetup #PyData