

WP3

casi scientifici

Luca Giommi
luca.giommi@cnafe.infn.it



- Laurea triennale e magistrale in Fisica all'Università di Bologna
- Dottorando in Data Science and Computation fino al 31/10/2022
- Assegnista al CNAF

- Temi di ricerca
 - Predizione della popolarità dei dataset di CMS
 - Soluzione di Predictive Maintenance nel centro di calcolo INFN-CNAF
 - Machine Learning as a Service for High Energy Physics
 - Manutenimento e sviluppo di nuove funzionalità per la PaaS di INFN Cloud

- Coinvolgimento in ML_INFN
 - Membro dell'unità operativa della sezione INFN di Bologna
 - Partecipazione agli hackathon



Raccolta delle esperienze esistenti nell'INFN nell'ambito del Machine Learning mediante la realizzazione di una knowledge base dove i singoli use case vengono documentati e corredati da esempi funzionanti

- Come entry point del progetto verso l'esterno e come catalogo degli use case usiamo Confluence <https://confluence.infn.it/display/MLINFN/Machine+Learning+Knowledge+Base>
- Knowledge base utile come connessione per collaborazioni (idee per attività di tesi?) sui vari use case

Table of Use cases

Name and Link	ML Technologies	Scientific Field	ML Tools	Comments
Btagging in CMS (templated version)	CNN, LSTM	High Energy Physics	Keras + Tensorflow	Realistic application
LHCb Masterclass, with Keras	DE, MLP	High Energy Physics	ROOT + Keras + TF	Introductory tutorial
MNIST in a C header	MLP		Keras	Free-styling tutorial
LUMIN: Lumin Unifies Many Improvements for Networks	CNN, RNN, GNN	High Energy Physics	PyTorch	Package use examples
INFERNO: Inference-Aware Neural Optimisation	NN	High Energy Physics	Keras + Tensorflow	Technique application example
An introduction to classification with CMS data	Fisher, BDT, MLP	High Energy Physics	Scikit-learn, TF2	Tutorials for Master Students
Virgo Autoencoder tutorial	Autoencoder	General Relativity	Python Keras	Tutorial for student
Distributed training of neural networks with Apache Spark	DNN	High Energy Physics	Spark + BigDL	Tutorial
FTS log analysis with NLP	NLP	High Energy Physics, Computing	Word2Vec + Rake + sklearn	
Image inpainting tutorial: how to digitally restore damaged images	CNN U-Net	Applied Physics	Keras + Sci-kit image, PIL, OpenCV, matplotlib	Tutorial
Signal/background discrimination for the VBF Higgs four lepton decay channel with the CMS experiment using Machine Learning classification techniques	ANNs, RF	High Energy Physics	Keras, TensorFlow, Scikit-learn	Tutorial
Explainability of a CNN classifier for breast density assessment	CNN	Medical Physics	Keras, Tensorflow	Tutorial
ML for smart caching	ML/RL	High Energy Physics, Computing, Cache	Keras, Tensorflow, sklearn	Demo, playground

- Informazioni da inserire
 - Autori con info sui contatti
 - Informazioni generali (livello difficoltà, campo scientifico, ...)
 - Tool, librerie e algoritmi ML utilizzati
 - Info sul/sui dataset da usare
 - Descrizione dello use case
 - Come riprodurre lo use case (fornendo codice)

Author(s)

Name	Institution	Mail Address	Social Contacts
Leonardo Giannini	INFN Sezione di Pisa / UCSD	leonardo.giannini@cern.ch	
Tommaso Boccali	INFN Sezione di Pisa	tommaso.boccali@pi.infn.it	Skype: tomboc73; Hangouts: tommaso.boccali@gmail.com

How to Obtain Support

Mail	tommaso.boccali@pi.infn.it , leonardo.giannini@cern.ch
Social	Skype: tomboc73
Jira	

General Information

ML/DL Technologies	LSTM, CNN
Science Fields	High Energy Physics
Difficulty	Low
Language	English
Type	fully annotated and runnable

📄 Presentation made on 📅 8-feb-2021 : <https://agenda.infn.it/event/25728/contributions/129752/attachments/78751/101917/go>

Software and Tools

Programming Language	Python
ML Toolset	Keras + Tensorflow
Additional libraries	uproot
Suggested Environments	INFN-Cloud VM, bare Linux Node, Google CoLab

Needed datasets

Data Creator	CMS Experiment
Data Type	Simulation

- Ampliare la collezione dei casi d'uso realistici che utilizzano tecniche di Machine Learning nelle varie linee di ricerca dell'Ente
 - La persona o gruppo che vuole proporre uno use case contatta me e Lucio e ne discutiamo
 - Presentare ad uno dei meeting settimanali di ML-INFN
 - Aggiungere lo use case alla knowledge base su Confluence
 - Utilizzo di un dataset pubblico

ML-INFN / Dashboard / ML-INFN Knowledge Base / Entry Point ML-INFN  BOZZA

Titolo della pagina

[Tabella dei contenuti](#)

Author(s)

Name	Institution	Mail Address	Social Contacts
@author1	INFN Division, Nat. Laboratory, Associated Group, Nat. Center, Consortia	author1 e-mail address	Skype: nickname; LinkedIn: nickname; Twitter: nickname; Hangouts: e-mail; Remove unused or add other contacts or write "N/A" in case none available
@authorX	& ... other institutions	authorX e-mail address	Skype: nickname; LinkedIn: nickname; Twitter: nickname; Hangouts: e-mail; Remove unused or add other contacts

How to Obtain Support

Mail	e-mails, mailing-lists
Social	Skype: nickname; LinkedIn: nickname; Twitter: nickname; Hangouts: e-mail; Remove unused or add other contacts or write "N/A" in case none available
Jira	add ticketing system endpoint in case you or your team are using one

General Information

ML/DL Technologies	add acronym of technologies used, ex: LTSM, CNN, ... others
Science Fields	ex: High Energy Physics, General Relativity, Computing
Difficulty	difficulty level of comprehension: low, medium, high
Language	Human language used to describe the Use Case
Type	Ex.: fully annotated, runnable, external resource, all of them

Software and Tools

Programming Language	Python, Java, C++,
ML Toolset	ML toolset or toolsets used. Write all that apply. Ex: