

FABIO PETRONI, PhD

Ricercatore in Informatica



📍 Londra, Regno Unito

@ fabio.petroni.1986@gmail.com

🌐 <http://www.fabiopetroni.com>

Sono un esperto ed appassionato di informatica. Le mie aree di interesse includono l'intelligenza artificiale, il machine learning, i sistemi distribuiti, il processamento automatico del linguaggio naturale, i sistemi di raccomandazione, i big data, le blockchain. Al momento sto lavorando su modelli basati su deep learning per l'estrazione di informazioni da grosse moli di dati nel reparto di Ricerca e Sviluppo di Thomson Reuters a Londra. In precedenza ho lavorato come Assistant Manager in KPMG. Ho conseguito un dottorato di ricerca in Ingegneria Informatica presso l'Università degli Studi di Roma "La Sapienza". Ho anche una laurea triennale e una laurea magistrale con lode in ingegneria informatica. Sono stato ricercatore presso l'Università di Mannheim e ho trascorso dei mesi presso l'Università di Groningen. Sono uno sviluppatore Spark certificato.

ESPERIENZA PROFESSIONALE

Research Scientist

Thomson Reuters

📅 Gennaio 2017 – Presente

📍 Londra



Machine Learning, Big Data, Natural Language Processing, Deep Neural Networks, Distributed Computation, Blockchain

TECNOLOGIE:

- TensorFlow, scikit-learn, Python
- Spark, Kafka, NLP4j, OpenNLP, Scala, Java
- Elasticsearch, AWS, GPU computation
- Smart Contracts, Solidity, Geth

Assistant Manager, Big Data Engineer

KPMG

📅 Aprile 2016 – Novembre 2016

📍 Londra



Soluzioni di apprendimento automatico in framework di Big Data

- Utilizzo estensivo di Hadoop e Spark (MLlib, GraphX, SQL, Streaming).
- Sviluppo di un ChatBot proprietario e della relativa base di conoscenza.
- Natural Language Processing - entity extraction, sentiment analysis.
- Metodi per la riduzione della dimensionalità (PCA) per la segmentazione dei clienti.
- Refactoring R/ Python scripts in Spark jobs - 1000X più veloce.
- Sviluppo di algoritmi di graph analytics su grandi grafi (milioni di nodi).
- Memorizzare e indicizzare dati su Elasticsearch, Hive and Accumulo.
- Analisi geospaziali con Spark + GeoMesa + Accumulo.

TECNOLOGIE:

- **Linguaggi di Programmazione:** Scala, Java, Python
- **Big Data Frameworks:** Apache Spark, Hadoop, Pig
- **Strumenti di Machine Learning:** TensorFlow, MLlib, Deeplearning4j
- **Databases:** Hive, Accumulo, MongoDB, Elasticsearch
- **Altro:** GeoMesa, GeoServer, Apache Kafka, AWS

CERTIFICAZIONI



SVILUPPATORE
APACHE SPARK
CERTIFICATO

COMPETENZE

Machine Learning	Artificial Intelligence	
Natural Language Processing	Big Data	
Recommender Systems	Data Science	
Algorithms	Blockchain	Data Mining
Social Media Analysis	Graph Analytics	
Distributed Systems	Computer Science	

SOFTWARE

Linguaggi di Programmazione:

Scala, Java, Python, Solidity, C, C++, SQL, Bash (Unix Shell), etc

Big Data Frameworks: Apache Spark, Hadoop, Spark, GraphLab

Strumenti di Machine Learning:

TensorFlow, MLlib, H2O, Deeplearning4j, Mahout, scikit-learn, nltk, libFM, NLP4j, etc

Databases: Hive, Accumulo, MongoDB, Elasticsearch, PostgreSQL, mysql, etc

Altro: Conoscenza di algoritmi e strutture dati, analisi di costo asintotico, multithreading, unit tests, stream processing, AWS.

PROGETTI OPEN SOURCE:



**libFM: Libreria
Factorization
Machine**

<https://github.com/fabiopetroni>

PREMI E RICONOSCIMENTI



FINANZIAMENTO PROGETTO RICERCA - 2015
SARE: Scalable Algorithms for Relation Extraction



FINALISTA TELECOM BIG DATA CHALLENGE - 2014
COLITA: Collaborative Interest-driven Targeted Advertising



LAUREATO ECCELLENTE SAPIENZA A.A. 2011-2012
Università degli Studi di Roma "La Sapienza"



PERCORSO DI ECCELLENZA - 2011
Università degli Studi di Roma "La Sapienza"

PUBBLICAZIONI SELEZIONATE

M. D. Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. Stanley, W. Quattrocioni:

The spreading of misinformation online.

In: Proceedings of the National Academy of Sciences (PNAS), 2016.

F. Petroni, L. Querzoni, K. Daudjee, S. Kamali and G. Iacoboni:

HDRF: Stream-Based Partitioning for Power-Law Graphs.

In: Proceedings of the 24th ACM International Conference on Conference on Information and Knowledge Management (CIKM), 2015.

F. Petroni, L. Del Corro and R. Gemulla:

CORE: Context-Aware Open Relation Extraction with Factorization Machines.

In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2015.

F. Petroni and L. Querzoni:

GASGD: Stochastic Gradient Descent for Distributed Asynchronous Matrix Completion via Graph Partitioning.

In: Proceedings of the 8th ACM Conference on Recommender Systems (RecSys), 2014.

Per la lista completa delle pubblicazioni: <http://www.fabiopetroni.com/publications>

EDUCAZIONE

DOTTORATO DI RICERCA in INGEGNERIA INFORMATICA



Università degli Studi di Roma "La Sapienza"

Novembre, 2012 - Febbraio, 2016

Roma, Italia

- **Argomenti: Big Data, Machine Learning, Data Mining**
- **Tesi: "Mining at scale with latent factor models for matrix completion"**

RISULTATI PRINCIPALI:

- Ideato un algoritmo di partizionamento di grafi che garantisce minima replicazione dei noti e bilanciamento del carico ottimale (attualmente è il migliore algoritmo di partizionamento grafi esistente).
- Migliorata la scalabilità degli algoritmi basati su fattorizzazione di matrici per i sistemi di raccomandazione (fino a 100 volte più veloce su cluster).
- Ideato modelli di completamento automatico per basi di conoscenza che incorporano informazioni contestuali per migliorare l'accuratezza.
- Analizzato la diffusione della disinformazione nei social networks.

EDUCAZIONE

Laurea Magistrale in Ingegneria Informatica

Università degli Studi di Roma "La Sapienza"

Ottobre, 2010 - Ottobre, 2012
Roma, Italia

- **Indirizzo: Architetture e Sistemi Distribuiti**
- **Tesi: "HSIENA: a hybrid publish/subscribe system."**
- **Voto Finale: 110 e Lode / 110**

MATERIE PRINCIPALI:

Sistemi Distribuiti, Computer and Networks Security, Sistemi Operativi, Big Data, Reti di Calcolatori, Ingegneria del Software, Database, Intelligenza Artificiale, Machine Learning, Capacity Planning e altri.

Laurea Triennale in Ingegneria Informatica

Università degli Studi di Roma "La Sapienza"

Ottobre, 2005 - Luglio, 2010
Roma, Italia

MATERIE PRINCIPALI:

Matematica, Statistica e Probabilità, Geometria, Fisica, Linguaggi di Programmazione, Scienze Informatiche, Sistemi Operativi, Sicurezza, Ricerca Operativa e altri.

ESPERIENZA ACCADEMICA

Research Scientist

UNIVERSITY OF MANNHEIM

Dicembre, 2014 - Maggio, 2015
Mannheim, Germania

Esplorato tecniche di machine learning basate su factorization machines per l'estrazione automatica di relazioni da linguaggio naturale (popolamento basi di conoscenza), una famiglia di algoritmi di natural language processing che mirano a predire nuovi fatti reali del mondo.

Programma Erasmus

RIJKSUNIVERSITEIT GRONINGEN

Febbraio, 2008 - Settembre, 2008
Groningen, Olanda

LINGUE

Italiano (madrelingua) ●●●●●

Inglese ●●●●●

Referenze disponibili su richiesta