

Diego Bettega

Data Scientist

Profilo

 +39 324 7728599

 diego.bettega.90@gmail.com

 [Via Consolatico Superiore n° 104, Busseto \(PR\), 43011, Italia](#)

 [Profilo LinkedIn](#)

 [Archivio GitHub](#)

Machine Learning & AI

Python Packages:

- Numpy, Pandas, Matplotlib, Seaborn
- Scikit Learn, Scipy, TensorFlow, Keras

Supervised Learning:

- Regularised Linear & Logistic models
- Naïve Bayes, KNN, Decision Tree, Kernel SVM
- Ensemble techniques: Bagging, Random Forest, Adaboost, XGBoost, Light GBM, Voting Classifier

Unsupervised Learning:

- K-means, Hierarchical Clustering

Dimensionality Reduction:

- PCA, Kernel PCA, LDA, SOMs, t-SNE, UMAP

Recommendation Systems:

- Market Basket Analysis, Collaborative Filtering
- SVD, RBM

Deep Learning:

- ANN, CNN, RNN-LSTM, GANs, AutoEncoders

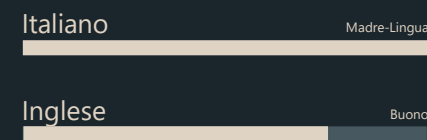
Reinforcement Learning (limited experience):

- UCB, Thompson Sampling
- Markov Decision Process

Competenze



Lingue



Esperienza Lavorativa

Mar 2018 -
Oggi

Snr Business Analyst / Data Scientist, EMEA

Kinesso (IPG) – Londra (UK) & Milano (IT)

- Pioniere e creatore di un modello globale per l'ottimizzazione automatica delle campagne di marketing, utilizzando modelli di machine learning (Python) e sviluppato in una React JS Web App. I costi sono stati ridotti in media del 30%.
- Automatizzato un processo di reporting raccogliendo dati da più origini di Google (ADH) e database di Amazon (AMC) utilizzando API e query SQL, analizzando i dati in Python e visualizzando i risultati in Tableau. Implementato un modello Bidirectional LSTM per path-to-conversion analisi e modelli ML per ottimizzare la frequenza di ad tra le campagne e generare insights.
- Progettato e implementato con successo una dashboard di Tableau con le informazioni sull'audience, utilizzando dati di prima parte, terze parti e programmatici analizzati utilizzando R e alteryx e queried utilizzando MySQL. Questo nuovo prodotto è attualmente utilizzato e venduto dal senior leadership team EMEA.
- Promosso tre volte nei primi 3 anni.

Lug 2016 -
Feb 2018

Data Analyst

Adloox – Londra (UK)

- Aumentato le entrate, revisione delle discrepanze e rilevamento di differenti tipi di inefficienze in base ai criteri di Brand Safety, Viewability e frodi.
- Automatizzato processi di reporting manuali che richiedevano molto tempo, potendo facilmente portare a errori (VBA Excel Macro).



Educazione

Giu 2019 -
Lug 2020

L'Università del Texas in Austin #2 in Analytics, #4 in AI, #7 in Machine Learning

➤ **PGP in AI and ML** (Voto: Eccellente)

Advanced statistic; supervised - unsupervised - reinforcement learning; featurization, model selection and tuning; recommendations systems; deep learning (ANN, CNN, RNN - LSTM); NLP; GANs; model deployment.

Vincitore di una competizione con 120 data scientist partecipanti. L'obiettivo era massimizzare l'accuratezza in un problema di classificazione supervisionato.

#4 posizione nella competizione con oltre 1000 data scientist partecipanti. Vinto un premio: corso in Mastering Big Data Analytics.

Giu 2020 -
Feb 2021

➤ **PGP in Cloud Computing**

Cloud Foundations; Specialization in AWS, Microsoft Azure and Google Cloud; Containers; Microservices; Big Data Management and Analytics on Cloud; Cloud Security & Migration; Private Cloud; Enterprise Cloud Solutions; Cloud-Native DevOps; On Prem DevOps.

Dic 2020 -
Present

➤ **Corso in Mastering Big Data Analytics**

Apache Hadoop; Map reduce; HDFS; YARN; Hive; Pyspark; Spark SQL; Spark MLlib; Spark streaming; Kafka.

Dic 2014

Ingegneria Meccanica

Laurea Triennale – Università di Parma, IT



Progetti

- **Serie Storiche:** Tecniche di ensemble e stateful LSTM sulla previsione del mercato FOREX.
- **NLP:** GloVe embedding e LSTM bidirezionale multistrato sul rilevamento del sarcasmo.
- **Rilevamento Facciale:** CNN (architettura UNET) per individuare la posizione del volto.
- **Riconoscimento Facciale:** VGG face, Triplet Loss, PCA e SVM per riconoscere i volti.
- **Classificazione delle immagini:** ANN per identificare le immagini contenenti numeri.
- **Sistema di Raccomandazione:** Consigli per i prodotti su un sito di e-commerce.
- **Fidelizzazione dei Clienti:** Stacking ML per identificare i clienti inclini all'abbandono.
- **Espleinabilità del Modello:** Grafico delle dipendenze parziali e SHAP per approfondimenti.
- **Soluzione di Comunicazione in Team:** Implementazione utilizzando Mattermost e AWS.
- **Servizi Gestiti su un Cloud Pubblico:** Processo aziendale automatizzato.
- **Web App:** Web App in ECS.
- **Concetto di Master'less arch:** Installa multi-node Cassandra cluster and inducil'errore.
- **Jenkins Server:** Gestisci in remoto le istanze EC2 e crea un server Jenkins