






Antonio Lanza


Data di nascita: 13/12/1996

CONTATTI

 antoniolanza1996@gmail.com

 (+39) 3405429172

 www.linkedin.com/in/lanzaantonio

 <https://github.com/antoniolanza1996> (GitHub)

ESPERIENZA LAVORATIVA

01/2021 – ATTUALE Rende, Italia

AI Research Scientist & ML Engineer ALTILIA s.r.l.

- Lavoro nel team R&D con l'obiettivo di costruire Altilia Intelligent Automation™ - una piattaforma no/low-code, IPaaS (Intelligent Process Automation as a Service) - che democratizza l'adozione dell'intelligenza artificiale, dell'iperautomazione e dell'intelligenza decisionale su larga scala nelle organizzazioni moderne di qualsiasi dimensione.
- Conduzione ricerche all'avanguardia nell'IA, inclusa l'esplorazione di nuovi algoritmi, tecniche e approcci di Intelligent Document Processing (IDP).
- Progettazione, implementazione e gestione di pipeline di machine learning pronte per la produzione per fornire soluzioni di IA affidabili e ad alte prestazioni che hanno un impatto diretto sull'efficienza dei processi aziendali.

06/2020 – 12/2020 Rende, Italia

AI Research Scientist Intern ALTILIA s.r.l.

Periodo di stage in cui sono stati studiati ed implementati modelli di Deep Learning Transformer-based, i quali rappresentano lo stato dell'arte nel campo del Natural Language Processing (NLP). In particolare, sono stati approfonditi modelli di Information Retrieval (IR) e Question Answering (QA) al fine di implementare un sistema di Open Domain Question Answering (ODQA).

ISTRUZIONE E FORMAZIONE

09/2021 – 09/2021

Abilitazione all'esercizio della professione di Ingegnere dell'Informazione (Sez. A dell'Albo) Università della Calabria

01/2020 – 05/2020 Dublino, Irlanda

Periodo Erasmus+ University College Dublin (UCD)

Semester GPA: 4.04/4.2

09/2018 – 12/2020 Rende, Italia

Laurea Magistrale in Ingegneria Informatica Università della Calabria

Percorso di studio: Big Data Analytics e Data Science

Titolo tesi: Tecniche di Deep Learning per l'Open Domain Question Answering

Voto finale: 110/110 con lode

09/2015 – 09/2018 Rende, Italia

Laurea Triennale in Ingegneria Informatica Università della Calabria

Titolo tesi: Uso degli Hidden Markov Models per l'interpretazione di business process logs di basso livello

Voto finale: 110/110 con lode

09/2010 – 07/2015 Castrovillari, Italia

Diploma di Liceo Scientifico Liceo Scientifico E. Mattei

COMPETENZE PROFESSIONALI

Competenze professionali

- **Generative AI e LLMs:**
 - Ampia esperienza con l'IA generativa in molteplici applicazioni:
 - **Retrieval-Augmented Generation (RAG) ed agenti di IA:** Progettato ed implementato diverse pipeline RAG ed agenti di IA, consentendo un recupero efficace delle informazioni e l'automazione di attività complesse.
 - **Prompt Engineering:** zero e few-shot prompt engineering
 - **Fine-Tuning of open-source LLMs:** Applicato tecniche PEFT (e.g. LoRA, QLoRA) per un adattamento efficiente degli LLMs a compiti specializzati.
 - *Librerie e framework:* PyTorch, Transformers, Datasets, PEFT e TRL di Hugging Face, LangChain, vLLM, OpenAI APIs.
- **Vector Databases e Information Retrieval:**
 - Ampia esperienza con tecniche di ricerca full-text (approcci sintattici/sparsi) e neurale (approcci semantici/densi) e con i database vettoriali per supportare il recupero di informazioni ad alte prestazioni basato sulla IA e la ricerca semantica in tempo reale.
 - *Librerie e framework:* Sentence Transformers di UKPLab, FARM e Haystack di deepset, Elasticsearch, Milvus, FAISS.
- **Altri task di Machine e Deep Learning:**
 - Ampia esperienza nello sviluppo e nell'implementazione di un'ampia gamma di modelli di ML/DL in vari domini:
 - **Natural Language Processing:** Named Entity Recognition, Question Answering, Multi-Class and Multi-Label Text Classification, Sentiment Analysis, Aspect-based Sentiment Analysis.
 - **Computer Vision:** Document Layout Analysis, Image Classification, Object Detection/Segmentation, Optical Character Recognition.
 - *Librerie e framework:* PyTorch, Transformers e Datasets di Hugging Face, Spacy, Scikit-learn, Weights & Biases, ClearML, Detectron2, MMCV e MMOCR di OpenMMLab.
- **DevOps e MLOps:**
 - Conoscenza delle pratiche Dev/MLOps per l'automazione ed il rilascio in produzione di modelli di ML/DL, con particolare attenzione al mantenimento di un'infrastruttura scalabile e resiliente.
 - Buona conoscenza nell'uso di Git, Shell scripting e Docker.
 - Conoscenza di base di Kubernetes e del framework Ray.
- **Linguaggi di programmazione:** Python e Java.

COMPETENZE LINGUISTICHE

LINGUA MADRE: italiano

Altre lingue:

inglese

Ascolto B2

Produzione orale B2

Lettura B2

Interazione orale B2

Scrittura B2

Livelli: A1 e A2: Livello elementare B1 e B2: Livello intermedio C1 e C2: Livello avanzato

CONFERENZE E SEMINARI

29/05/2024 – 30/05/2024 Ital-IA 2024 - Napoli, Italy

Evaluating Retrieval-Augmented Generation for Question Answering with Large Language Models

Abstract: We present a comprehensive framework for evaluating retrieval-augmented generation (RAG) systems designed for questionanswering tasks using large language models (LLMs). The proposed framework integrates document ingestion, information retrieval, answer generation, and evaluation phases. Both ground truth-based and reference-free evaluation metrics are implemented to provide a multi-faceted assessment approach. Through experiments across diverse datasets like NarrativeQA and a proprietary financial dataset (FinAM-it), the reliability of existing metrics is investigated by comparing them against rigorous human evaluations. The results demonstrate that ground truth-based metrics such as BEM and RAGAS Answer Correctness exhibit a moderately strong correlation with human judgments. However, reference-free metrics still struggle to capture nuances in

answer quality without predefined correct responses accurately. An in-depth analysis of Spearman correlation coefficients sheds light on the interrelationships and relative effectiveness of various evaluation approaches across multiple domains. While highlighting the current limitations of reference-free methodologies, the study underscores the need for more sophisticated techniques to better approximate human perception of answer relevance and correctness. Overall, this research contributes to ongoing efforts in developing reliable evaluation frameworks for RAG systems, paving the way for advancements in natural language processing and the realization of highly accurate and human-like AI systems.

Link <https://ceur-ws.org/Vol-3762/495.pdf>

29/05/2023 – 31/05/2023 Ital-IA 2023 - Pisa, Italy

● **Building a Platform for Intelligent Document Processing: Opportunities and Challenges**

Abstract: Companies of any size and industry still struggle in automatic business processes where human cognitive and contextualization capabilities are required to read and understand complex documents. Ongoing progress in the fields of Computer Vision and Natural Language Processing, where (large) language models are becoming increasingly and freely available, have made possible to create a new generation of Intelligent Document Processing technologies that allow automatically analyzing and understanding both documents layout and contents. In this paper we present an Intelligent Document Processing platform that makes use of hybrid AI techniques to allow document reading comprehension by means of a combination of Document Layout Analysis and recognition, table recognition and detection, context free grammars, and question answering techniques. Such a technology combines also no-code principles with high performance computing based on micro-services to streamline the execution of tasks such as document and text classification, document segmentation, entity extraction, sentiment analysis, question answering, and more.

Link <https://ceur-ws.org/Vol-3486/91.pdf>

24/04/2023 – 26/04/2023 ICEIS 2023 - Prague, Czech Republic

● **ESG Data Collection with Adaptive AI**

Abstract: The European Commission defines the sustainable finance as the process of taking Environmental, Social and Governance (ESG) considerations into account when making investment decisions, leading to more long-term investments in sustainable economic activities and projects. Banks, and other financial institutions, are increasingly incorporating data about ESG performances, with particular reference to risks posed by climate change, into their credit and investment portfolios evaluation methods. However, collecting the data related to ESG performances of corporate and businesses is still a difficult task. There exist no single source from which we can extract all the data. Furthermore, most important ESG data is in unstructured format, hence collecting it poses many technological and methodological challenges. In this paper we propose a method that addresses the ESG data collection problem based on AI-based approaches. We also present the implementation of the proposed method and discuss some experiments carried out on real world documents.

Link <https://www.scitepress.org/PublicationsDetail.aspx?ID=sCYU36yksi=&t=1>

08/11/2021 – 10/11/2021 Online Conference

● **Codemotion Online Tech Conference 2021 - Italian Edition | Autumn**

ONORIFICENZE E RICONOSCIMENTI

06/2020 Dipartimento DIMES, Università della Calabria

● **Percorso di Eccellenza**

Vincitore di una borsa di studio indetta dal dipartimento DIMES.

Durante questo percorso è stato approfondito il framework Apache Storm realizzando un'applicazione che analizza tweet pubblicati su Twitter in tempo reale e fornisce, mediante opportuni grafici, delle statistiche sugli hashtag e sulla provenienza degli utenti. Inoltre, è possibile anche eseguire task di sentiment analysis e spam detection.

24/05/2016 PRAXIS MMT Corp, per conto di Business Talents, Milano (MI)

● **Finalista del programma di Impresa Simulata BT – Edizione 2015**

09/04/2015 PRAXIS MMT Corp, per conto di Young Business Talents, Milano (MI)

● **Vincitore del programma di Impresa Simulata YBT – Edizione 2014**

21/03/2014 PRAXIS MMT Corp, per conto di Young Business Talents, Milano (MI)

● **Finalista del programma di Impresa Simulata YBT – Edizione 2013**