

CURRICULUM VITAE

Riccardo Cantini

Contents

1	Informazioni generali	2
2	Posizioni ricoperte	2
3	Istruzione e formazione	2
4	Attività di ricerca	3
4.1	Metodologie e algoritmi per l'analisi dei Big Social Data	3
4.2	Tecniche per ottimizzare l'esecuzione di applicazioni data-intensive	4
4.3	Linee di ricerca correnti e future	4
5	Elenco delle pubblicazioni	4
6	Partecipazione in qualità di relatore a conferenze di interesse internazionale	5
7	Partecipazione a progetti	6
8	Attività editoriale	6
9	Attività didattica	6
10	Supervisione di tesi di laurea	7

1 Informazioni generali

CODICE FISCALE	DATA DI NASCITA	LUOGO DI NASCITA	NAZIONALITÀ
CNTRCR93T30C352C	30 dicembre, 1993	Catanzaro (CZ), Italia	Italiana
INDIRIZZO	TELEFONO	E-MAIL	
Lavoro: via Pietro Bucci 41C, 87036 Rende (CS)	+39 3272697509	riccardo.cantini@unical.it	
Casa: via Teano n.13, 88100 Catanzaro (CZ)			

2 Posizioni ricoperte

- Da luglio 2023, *Ricercatore a Tempo Determinato di Tipo A*, SSD ING-INF/05, presso DIMES, Università della Calabria, Rende (CS).
- Da marzo 2023 a luglio 2023, *borsista di ricerca* presso DIMES, Università della Calabria, Rende (CS), nell'ambito del progetto *ASPIDE*.
Titolo della ricerca: *Tecniche di topic detection and tracking da dati provenienti dai social media*.
Responsabile scientifico: prof. Fabrizio Marozzo.
- Da aprile 2021 a luglio 2022, *visiting researcher* presso il dipartimento di Computer Science, Workflows and Distributed Computing Group, Barcelona Supercomputing Center (BSC), Spagna.
Responsabile scientifico: prof. Rosa M. Badia.
- Da novembre 2019 a marzo 2023, *dottorando in Information and Communication Technologies* presso DIMES, Università della Calabria, Rende (CS).
- Da luglio 2019 a ottobre 2019, *assegnista di ricerca* presso DIMES, Università della Calabria, Rende (CS), nell'ambito del progetto *ASPIDE*.
Titolo della ricerca: *Tecniche in-memory per l'esecuzione efficiente di applicazioni data-intensive su architetture Exascale*.
Responsabile scientifico: prof. Paolo Trunfio.
- Da luglio 2018 a giugno 2019, *collaboratore di ricerca* presso DIMES, Università della Calabria, Rende (CS), nell'ambito del progetto *Smart Macingo*.
Titolo della ricerca: *Tecniche di data analysis per la stima dei prezzi nei servizi di trasporto*.
Responsabile scientifico: prof. Paolo Trunfio.

3 Istruzione e formazione

- Dottorato di ricerca europeo in Information and Communication Technology, DIMES, Università della Calabria, Rende (CS).
Titolo tesi: *Distributed Big Social Data Analysis: Advanced Techniques and Execution Strategies*.
Supervisore: prof. Paolo Trunfio.
Co-supervisore: prof. Fabrizio Marozzo.
Votazione: Eccellente
Data conseguimento: 07/03/2023.

- Partecipazione alla 1st International School on Internet of Things and Edge AI: Computing, Communications and Systems, Falerna (CZ).

Ente organizzatore: DIMES, Università della Calabria, con il patrocinio della IEEE Italy section e del Gruppo di Ingegneria Informatica (GII).

Data: 08-12/09/2022.

- Laurea Magistrale in Ingegneria Informatica, Università della Calabria, Rende (CS)
Titolo tesi: *Analisi della polarizzazione politica degli utenti di Twitter mediante reti neurali*.
Relatori: prof. Paolo Trunfio, prof. Fabrizio Marozzo.
Votazione: 110/110 con Lode.
Data conseguimento: 12/04/2019.
- Laurea Triennale in Ingegneria Informatica, Università della Calabria, Rende (CS).
Titolo tesi: *Uso del framework ELKI per la scoperta di regioni d'interesse tramite tecniche di clustering*.
Relatori: prof. Paolo Trunfio, prof. Fabrizio Marozzo.

4 Attività di ricerca

L'ambito principale dell'attività di ricerca di Riccardo Cantini riguarda l'analisi dei Big Data, in particolare quelli prodotti sulle principali piattaforme di social media, tramite tecniche di machine e deep learning e framework per il calcolo parallelo e distribuito. L'obiettivo della ricerca è duplice: da un lato, mira allo sviluppo e alla valutazione di metodologie e algoritmi innovativi per l'analisi dei Big Social Data, da cui estrarre informazioni preziose sulle dinamiche, il comportamento e l'opinione degli utenti; dall'altro, è volta alla progettazione e all'implementazione di nuove strategie e tecniche per supportare l'analisi efficiente di tali dati in ambienti distribuiti e infrastrutture di calcolo ad alte prestazioni.

4.1 Metodologie e algoritmi per l'analisi dei Big Social Data

L'attività di ricerca si è focalizzata prevalentemente sull'analisi dei Big Social Data a carattere politico, prodotti dagli utenti social durante campagne elettorali e referendum. L'obiettivo principale è quello di delineare un profilo dettagliato degli utenti social politicamente polarizzati, identificando i principali temi su cui verte la conversazione online e indagando le loro opinioni e sentimenti. In tal modo è possibile modellare la loro percezione di fatti ed eventi del mondo reale, fornendo un approccio data-driven ad una comprensione approfondita dei fenomeni politici.

In particolare, in [1], [2] è stata proposta *IOM-NN (Iterative Opinion Mining using Neural Networks)*, una metodologia semi-supervisionata che sfrutta un processo iterativo di pseudo-labeling basato su reti neurali per la stima della polarizzazione politica degli utenti social. Tale metodologia è stata estesa in [3], integrando l'analisi della polarizzazione politica con tecniche di topic detection ed emotion analysis, con l'obiettivo di modellare il sostegno politico attraverso un ampio spettro di emozioni, delineando una rappresentazione accurata dell'evento politico d'interesse attraverso un workflow d'analisi unificato.

Al fine di far fronte ai problemi legati alla dinamicità e all'affidabilità dei dati coinvolti nei processi di analisi della polarizzazione politica degli utenti social, in [4] è stata proposta *TIMBRE (Time-aware opInion Mining via Bot REMoval)*, una metodologia di opinion mining che stima le intenzioni di voto degli utenti attraverso un processo hashtag-based. Tale metodologia sfrutta informazioni temporali durante il processo di classificazione, risultando dunque robusta rispetto alle oscillazioni temporali della polarizzazione politica degli utenti social. Inoltre, i contenuti pubblicati dai social bot vengono identificati e filtrati, valutandone l'impatto tramite processi di diffusione dell'informazione.

Per ottenere una rappresentazione completa della conversazione sui social media, sono state sviluppate tecniche di topic detection e tracking, finalizzate all'identificazione degli argomenti a sfondo politico e sociale alla base della conversazione online e allo studio della loro evoluzione nel tempo [5]. L'analisi

della polarizzazione politica degli utenti social è stata inoltre integrata all'interno di modelli di diffusione dell'informazione e tecniche di massimizzazione dell'influenza. In particolare, in [6] viene proposto un algoritmo di influence maximization bio-ispirato, volto all'identificazione degli utenti più influenti per ciascuna fazione e delle principali strategie di diffusione dell'informazione adottate durante la campagna politica.

Un'ulteriore area di ricerca è costituita dall'hashtag recommendation, in cui una nuova metodologia, basata su tecniche avanzate di elaborazione del linguaggio naturale (NLP), è stata proposta per determinare gli hashtag più appropriati per un determinato post [7]. Tale tecnica, denominata *HASHET (HAShtag recommendation using Sentence-to-Hashtag Embedding Translation)*, si basa sul mapping tra il contenuto semantico di un post e la rappresentazione latente dei suoi hashtag. Il processo di raccomandazione si basa poi sul concetto di *località* nello spazio latente degli hashtag, il che rende la metodologia capace di sfruttare a pieno la struttura di clustering sottesa a tale spazio, la quale riflette i principali topic di discussione e le relazioni semantiche tra hashtag.

4.2 Tecniche per ottimizzare l'esecuzione di applicazioni data-intensive

Nonostante l'analisi dei Big Social Data offra numerose opportunità in diversi ambiti, il loro volume e la loro velocità mettono continuamente a dura prova i sistemi odierni, in termini di storage, capacità di elaborazione e analisi [8]. In virtù di ciò, l'attività di ricerca si è anche focalizzata sulla progettazione e l'implementazione di nuove tecniche e strategie ad-hoc volte a consentire l'esecuzione efficiente di applicazioni data-intensive su architetture parallele e distribuite. Nello specifico, in [9] è stata presentata *IWM (Intelligent In-memory Workflow Manager)*, una metodologia che ottimizza l'esecuzione di workflow data-intensive bilanciando il livello di parallelismo e l'uso della memoria principale, al fine di minimizzare il data spilling su disco. IWM usa un modello di machine learning per predire l'occupazione di memoria ed il tempo di esecuzione dei task di un workflow, sfruttando tale informazione per la generazione di uno scheduling, sulla base di una soluzione euristica al problema del bin packing. Infine, in [10] viene proposto un approccio basato su machine learning per l'identificazione di un partizionamento adeguato che possa supportare l'esecuzione di applicazioni data-parallel in ambienti di calcolo parallelo e distribuito e infrastrutture ad alte prestazioni.

4.3 Linee di ricerca correnti e future

Oltre alle linee di ricerca precedentemente menzionate, altri ambiti vengono correntemente esplorati tra cui l'utilizzo dei Large Language Models, tecniche di meta learning, modelli semi-supervisionati basati su curriculum learning, tecniche neurali di topic detection [11], topic tracking e topic-oriented misinformation detection [12], e l'analisi di Big Social Data tramite tecniche di edge AI e modelli time-adaptive.

5 Elenco delle pubblicazioni

Riviste

- [2] L. Belcastro, R. Cantini, F. Marozzo, D. Talia, and P. Trunfio, "Learning political polarization on social media using neural networks," *IEEE Access*, vol. 8, pp. 47 177–47 187, 2020. DOI: 10.1109/ACCESS.2020.2978950.
- [3] L. Belcastro, F. Branda, R. Cantini, F. Marozzo, D. Talia, and P. Trunfio, "Analyzing voter behavior on social media during the 2020 us presidential election campaign," *Social Network Analysis and Mining*, vol. 12, no. 1, p. 83, 2022. DOI: 10.1007/s13278-022-00913-9.
- [4] R. Cantini, F. Marozzo, D. Talia, and P. Trunfio, "Analyzing political polarization on social media by deleting bot spamming," *Big Data and Cognitive Computing*, vol. 6, no. 1, 2022. DOI: 10.3390/bdcc6010003.
- [6] R. Cantini, F. Marozzo, S. Mazza, D. Talia, and P. Trunfio, "A weighted artificial bee colony algorithm for influence maximization," *Online Social Networks and Media*, vol. 26, p. 100 167, 2021. DOI: 10.1016/j.osnem.2021.100167.

- [7] R. Cantini, F. Marozzo, G. Bruno, and P. Trunfio, “Learning sentence-to-hashtags semantic mapping for hashtag recommendation on microblogs,” *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 16, no. 2, pp. 1–26, 2021. DOI: 10.1145/3466876.
- [8] L. Belcastro, R. Cantini, F. Marozzo, A. Orsino, D. Talia, and P. Trunfio, “Programming big data analysis: Principles and solutions,” *Journal of Big Data*, vol. 9, no. 1, pp. 1–50, 2022. DOI: 10.1186/s40537-021-00555-2.
- [9] R. Cantini, F. Marozzo, A. Orsino, D. Talia, and P. Trunfio, “Exploiting machine learning for improving in-memory execution of data-intensive workflows on parallel machines,” *Future Internet*, vol. 13, no. 5, 2021. DOI: 10.3390/fi13050121.
- [13] L. Belcastro, R. Cantini, and F. Marozzo, “Knowledge discovery from large amounts of social media data,” *Applied Sciences*, vol. 12, no. 3, 2022. DOI: 10.3390/app12031209.

Conferenze

- [1] L. Belcastro, R. Cantini, F. Marozzo, D. Talia, and P. Trunfio, “Discovering political polarization on social media: A case study,” in *2019 15th International Conference on Semantics, Knowledge and Grids (SKG)*, IEEE, 2019, pp. 182–189. DOI: 10.1109/SKG49510.2019.00038.
- [5] R. Cantini and F. Marozzo, “Topic detection and tracking in social media platforms,” in *Pervasive Knowledge and Collective Intelligence on Web and Social Media*, C. Comito and D. Talia, Eds., Cham: Springer Nature Switzerland, 2023, pp. 41–56. DOI: 10.1007/978-3-031-31469-8_3.
- [14] R. Cantini, F. Marozzo, A. Orsino, M. Passarelli, and P. Trunfio, “A visual tool for reducing returns in e-commerce platforms,” in *2021 IEEE 6th International Forum on Research and Technology for Society and Industry (RTSI)*, IEEE, 2021, pp. 474–479. DOI: 10.1109/RTSI50628.2021.9597230.

Lavori sottomessi

- [10] R. Cantini, F. Marozzo, A. Orsino, *et al.*, “Block size estimation for data partitioning in hpc applications using machine learning techniques,” Submitted to *Journal of Big Data*, 2022.
- [11] R. Cantini, F. Marozzo, D. Talia, and P. Trunfio, “Discovering ukraine’s war topics on twitter by clustering latent hashtag representations,” in Submitted to *2023 26th International Conference on Discovery Science (DS23)*, 2023.
- [12] R. Cantini, C. Cosentino, F. Marozzo, I. Kilanioti, D. Talia, and P. Trunfio, “Unmasking covid-19 false information on twitter: A topic-based approach with bert,” in Submitted to *2023 26th International Conference on Discovery Science (DS23)*, 2023.

6 Partecipazione in qualità di relatore a conferenze di interesse internazionale

Riccardo Cantini è stato relatore nei seguenti congressi internazionali e per la presentazione dei lavori indicati:

- *EAI International Conference on Pervasive knowledge and collective intelligence on Web and Social Media* (EAI PerSoM 2022) Messina, Italia, 17-18 novembre 2022, presentando il seguente lavoro: *Topic detection and tracking in social media platforms* [5].
- *IEEE 6th International Forum on Research and Technology for Society and Industry* (RTSI 2021), 6-9 settembre 2021, presentando il seguente lavoro: *A visual tool for reducing returns in e-commerce platforms* [14].

7 Partecipazione a progetti

- *eFlows4HPC: enabling dynamic and Intelligent workflows in the future EuroHPC ecosystem*

Scopo della ricerca: uso di tecniche di machine learning per l'ottimizzazione del data partitioning, al fine di supportare l'esecuzione efficiente di workflow data-intensive in ambienti HPC.

- *ASPIDE: exAScale ProgramIng models for extreme Data procEssing*

Scopo della ricerca: sviluppo di tecniche in-memory per l'esecuzione efficiente di applicazioni data-intensive su architetture Exascale.

- *Smart Macingo*

Scopo della ricerca: definizione ed implementazione di tecniche di data mining per la stima del prezzo dei servizi di trasporto.

8 Attività editoriale

Partecipazione a comitati di programma di conferenze

Riccardo Cantini è membro del Program Committee di ADMA, *Advanced Data Mining and Applications* 2023 (Core rank: B).

Revisioni per riviste

- *ACM Computing Surveys*: 2022
- *Transactions on Cloud Computing*: 2021
- *Transactions on Parallel and Distributed Systems*: 2022
- *Future Generation Computer Systems*: 2022, 2023
- *IEEE Access*: 2020, 2021
- *Computer*: 2023
- *Social Network Analysis and Mining*: 2022

Revisioni per conferenze

- *International Conference on Advanced Data Mining and Applications (ADMA)*: 2023
- *International European Conference on Parallel and Distributed Computing (EuroPar)*: 2023
- *IEEE International Conference on Machine Learning and Applications (ICMLA)*: 2021, 2022
- *IEEE International Conference on Big Data (IEEE BigData)*: 2019, 2020, 2021, 2022

9 Attività didattica

- **Sistemi Operativi.** Laurea Triennale in Ingegneria Informatica, DIMES, Università della Calabria, Rende (CS).

Argomenti trattati: principali costrutti per la modellazione e l'implementazione di applicazioni multi-thread, problematiche riguardanti la sincronizzazione e l'accesso a variabili condivise, meccanismi di concorrenza di Java, in particolare Semafori e Monitor.

A.A. 2022-2023: 27 ore di esercitazione; A.A. 2021-2022: 17 ore di esercitazione.

- **Distributed Systems and Cloud/Edge Computing for IoT.** Laurea Magistrale in Computer Engineering for the Internet of Things, DIMES, Università della Calabria, Rende (CS).

Argomenti trattati: principali concetti alla base dell'Edge e Fog Computing in relazione alle architetture Cloud e Internet of Things, con particolare attenzione ai principali strumenti e framework per la modellazione, la simulazione e l'implementazione di applicazioni context-aware su larga scala.

A.A. 2021-2022: 23 ore di esercitazione; A.A. 2020-2021: 23 ore di esercitazione; A.A. 2019-2020: 23 ore di esercitazione.

10 Supervisione di tesi di laurea

Riccardo Cantini è stato relatore di un totale di **30 tesisti** presso il DIMES, Università della Calabria, di cui **21 triennali** e **9 magistrali**, tra il 2019 e il 2023.

I principali argomenti delle tesi supervisionate includono: social media and big data analysis, machine learning, deep learning, natural language processing, sentiment analysis, opinion mining, edge/fog computing, high-performance data analytics, curriculum learning, semi-supervised learning, hashtag recommendation, topic detection and tracking.

Autorizzo al trattamento dei dati personali secondo quanto previsto dal regolamento UE GDPR 679/2016.

Dichiaro che le informazioni riportate nel presente Curriculum Vitae sono esatte e veritiere. Dichiaro di essere consapevole di quanto comporta l'affermazione della veridicità di quanto sopra rappresentato e di essere a conoscenza delle sanzioni penali di cui all'art. 76 del D.P.R. 28.12.2000 n.445 "Testo unico delle disposizioni legislative e regolamentari in materia di documentazione amministrativa" ed in particolare di quanto previsto dall'art. 495 del Codice Penale in caso di dichiarazioni mendaci o di false attestazioni. Quanto sopra viene presentato sotto forma di autocertificazione (dichiarazione sostitutiva di certificazione e dichiarazione sostitutiva di atto di notorietà) ai sensi degli artt. 19, 46 e 47 del DPR n. 445/2000.

Data

09/07/2023

Firma

RICCARDO CANTINI