



UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II

itee_{PhD}
information technology
electrical engineering



Valerio La Gatta

eXplainable Artificial Intelligence for multimodal social media analysis

Tutor: Prof. Vincenzo Moscato

Cycle: XXXVI

Year: First

My background

- MSc degree in Computer Engineering from University of Naples Federico II
- Research group: Pattern Analysis and Intelligence Computation for mUltimedia System (PICUS)
- PhD start date: 01/11/2020
- Scholarship type: Unina

Research field of interest

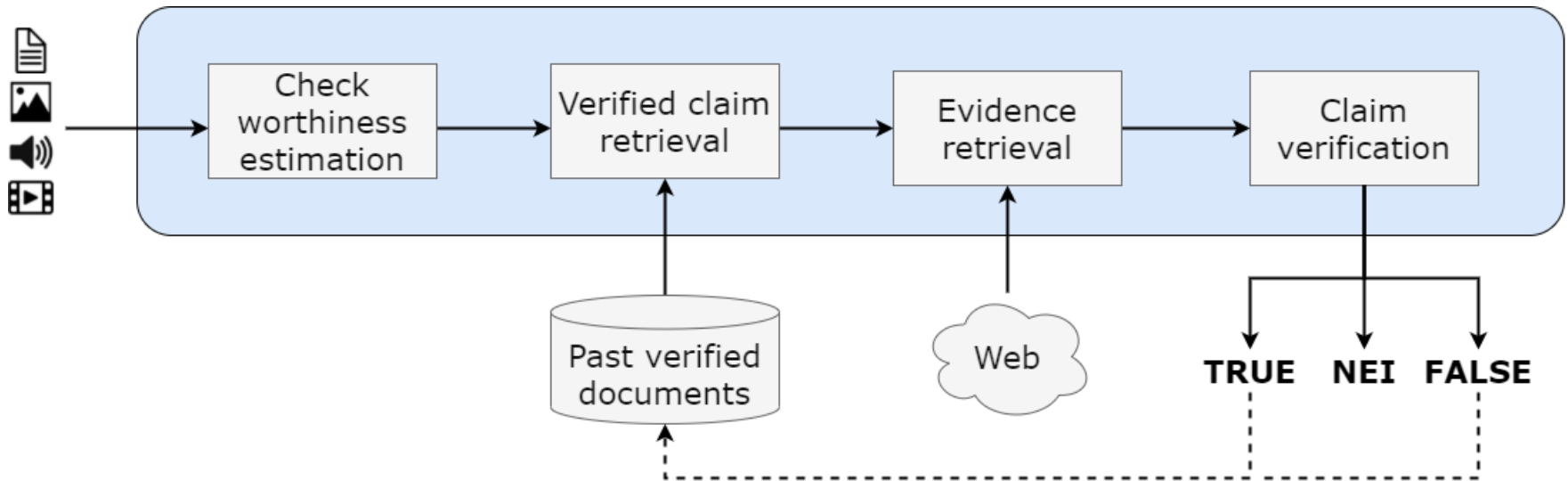
- Automatic fact-checking and fake news detection in social media
 - Support manual fact-checkers work and automate the verification process
 - The role of multimodal information (e.g. text, image, audio) within the fact-checking pipeline
 - Neural fake news
 - Cross-domain and time-varying fake news detection
- eXplainable Artificial Intelligence (XAI)
 - “Enable human users to understand, appropriately trust, and effectively manage the emerging generation of intelligent systems”
 - Explain why a certain model makes certain predictions (local vs global approaches, model-agnostic vs model-specific techniques)
 - Analyze models and data biases leading to unfair decisions

Summary of study activities

- Ad hoc PhD courses:
 - Statistical data analysis for science and engineering research
 - Data science for patient records analysis
 - Strategic Orientation for STEM research & writing
 - Scientific Programming and Visualization with Python
- Courses attended borrowed from MSc curricula
 - Natural language processing, prof. Cutugno
- PhD school:
 - AIRO PhD School 2021 and 5th AIRO-Young Workshop

Research activity: fact-checking

- *Context:* the fact-checking pipeline

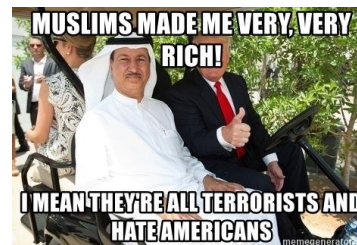


Research activity: fact-checking

- *Problem*: verified claim retrieval, a.k.a. detecting previously fact-checked information
 - Increase manual fact-checkers productivity
 - Improve automatic fact-checking
- *Objective*: ranking a list of verified documents according to the relevance with an input claim.
- *Methodology*: a two-stages retriever-reranker pipeline
 - Benchmark model neural approaches and standard probabilistic ones

Research activity: fact-checking

- *Challenges & Future works*
 - The role of multimodal data (e.g. text, images, audio) within the detection process
 - The impact of time on veracity classifiers
 - The connection between fake news detection and hate speech diffusion



Research activity: XAI

- *Context*: local model-agnostic and feature importance-based techniques
- *Problem*: it is not clear how explanations could be applied to new instances
 - Define the *explanation locality* through rules and provide *counterfactuals* to the prediction
- *Objective*: enrich the explanation with information about how the prediction would change when features changes
- *Methodology*: incorporate into the explanation process the global knowledge learned by the model
 - Describe global knowledge in terms of *pivots*, i.e. representative points of regions where the model exhibits different behaviors
 - Find the supporting (opposing) direction to the prediction, i.e. the changes that would make its probability increase (decrease)

Research activity: XAI

- *Challenges & Future works:*
 - Analyzing XAI domain-specific requirements (e.g. malware detection, fact-checking, social network analysis)
 - Extending the proposed methodology to deal with image classification tasks

```

InternetOpenA InternetOpenUrlA InternetCloseHandle
OpenSCManagerA CreateServiceA StartServiceA
LdrGetDllHandle LdrGetProcedureAddress FindResourceA
LoadResource SizeofResource NtOpenFile NtClose
NtCreateFile NtWriteFile NtClose CreateProcessInternalW
LdrGetDllHandle NtTerminateProcess NtClose
NtUnmapViewOfSection NtClose RegCloseKey NtClose
NtUnmapViewOfSection NtClose RegCloseKey NtClose
LdrUnloadDll NtOpenKey NtQueryValueKey NtClose
LdrUnloadDll NtClose NtTerminateProcess
    
```

Source: <https://dl.acm.org/doi/10.1145/3292500.3330935>

Sarah Palin Calls To Boycott Mall Of America Because “Santa Was Always White In The Bible”

... that comes to mind, many were highly offended. Three years after Fox’s Megyn Kelly definitively explained to America that both Jesus Christ and Santa Claus were white men, Mall of America dismissed her advice and hired Larry Jefferson, a retired ...

... need to run it into the ground, so that they never ever come up with such an offensive and sacrilegious idea again,” Palin added. “The Holy Book always said Santa Claus was white and any kind of deviation from that, regardless of its magnitude, is a sin. And we need to make an example out of Mall of America. If we

User 1 9 Dec 2016
Gee, did Santa and Jesus hang out and pound down a few beers together?

User 2 9 Dec 2016
St. Nicholas was white? Really?? Lol

User 3 9 Dec 2016
FYI, this is false.

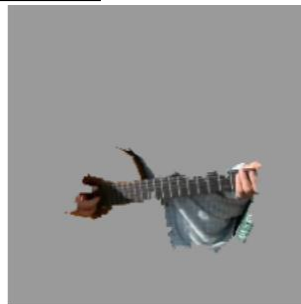
User 4 9 Dec 2016
OMG, Santa is in the Bible

User 5 9 Dec 2016
I wanted that one. I really did.

Source: <https://arxiv.org/abs/1602.04938>



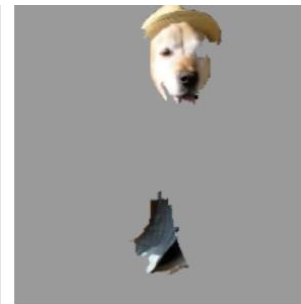
(a) Original Image



(b) Explaining *Electric guitar*



(c) Explaining *Acoustic guitar*



(d) Explaining *Labrador*

Valerio La Gatta

Products

[P1]	V. La Gatta , V. Moscato, M. Postiglione, G. Sperli; “An Epidemiological Neural network exploiting Dynamic Graph Structured Data applied to the COVID-19 outbreak”; IEEE Transactions on Big Data, IEEE TBD; published; 2020
[P2]	V. La Gatta , V. Moscato, M. Postiglione, G. Sperli; “CASTLE: Cluster-Aided Space Transformation for Local Explanations”; Expert Systems with Applications, ESWA; published; 2021
[P3]	V. La Gatta , V. Moscato, M. Postiglione, G. Sperli; “PASTLE: Pivot-Aided Space Transformation for Local Explanations”; Pattern Recognition Letters, PRL; published; 2021
[P4]	T. Chakraborty, V. La Gatta , V. Moscato, G. Sperli; “Information retrieval algorithms and neural ranking models to detect previously fact-checked information”; Knowledge-Based Systems, KBS; submitted
[P5]	A. Galli, V. La Gatta , V. Moscato, M. Postiglione, G. Sperli; “Interpretability in AI-based Behavioral Malware Detection Systems”, IEEE Transactions on Dependable and Secure Computing, IEEE TDSC; submitted

Tutorship

- Co-supervisor of five master theses in Computer Engineering regarding eXplainable AI, music recommender systems and automatic fact-checking
- Weekly two hours of teaching activities regarding practical lectures/seminars during the courses of “Hardware and Software for Big Data” (Master Degree in Data Science) and “Machine Learning and Big Data per la salute” (Master Degree in Biomedical Engineering)

Thank you for the attention!