Giovanna Nicora

Curriculum Vitae

Personal Data

Date and place of birth	January 9, 1992, Voghera (PV), Italy
Citizenship	Italy
Email	giovanna.nicora@unipv.it
Web site	`http://lab-bioinfo.unipv.it/index.php/people/65-nicorag`
Github	`https://github.com/GiovannaNicora`
Publications	Google Scholar, ORCID ID

Current position

From July 2023 **Researcher (RTDa)**, *University of Pavia*, Pavia (Italy).

Experience

December 2021 - July 2023 **Senior Bioinformatician, Artificial Intelligence Scientist**, *enGenome S.r.l*, Pavia (Italy).

December 2020-November 2021 **Post-Doc Research Fellow**, *Università degli Studi di Pavia*, Department of Electrical, Computer and Biomedical Engineering , Pavia (Italy).

Advisor: *Prof. Riccardo Bellazzi*

October 2017-November 2020 **PhD Student in Health Technology, Bioengineering and Bioinformatics**, *Università degli Studi di Pavia*, Department of Electrical, Computer and Biomedical Engineering, Pavia (Italy).

Advisor: *Prof. Riccardo Bellazzi*

February 2019–July 2019 **Visiting PhD student**, *Bioinformatics Laboratory, Faculty of Computer and Information Science*, The University of Ljubljana, Ljubljana (Slovenia).

Supervisor: *Prof. Blaž Zupan*

January 2017–September 2017 **Research Scholarship**, *Università degli Studi di Pavia*, Department of Electrical, Computer and Biomedical Engineering, Pavia (Italy).

Advisor: *Prof. Riccardo Bellazzi*

Education

2017–2020 **PhD in Bioengineering, Bioinformatics and Health Technologies**, *Università degli Studi di Pavia*, Pavia (Italy).

Duration of the degree: *3 years*

Thesis title: *Artificial Intelligence Strategies For Genomic Variant Interpretation in Hematological Cancer*

2014–2016 **Master's in Bioengineering**, *Università degli Studi di Pavia*, Pavia (Italy).

Duration of the degree: *2 years*

Score: *110 out of 110*

Thesis title: *Development of a rule based expert system for the automatic interpretation of genomic variant according to ACMG guidelines*

2011–2014 **Bachelor in Bioengineering**, *Università degli Studi di Pavia*, Pavia (Italy).

Duration of the degree: *3 years*

Score: *107 out of 110*

2006–2011 **Scientific high school diploma**, *Istituto di Istruzione Superiore Galileo Galilei*, Voghera (Italy).

Duration: *5 years*

Score: *97 out of 100*

Technical skills

Artificial Intelligence, Machine Learning and and Data analysis

- Strong experience with the development of Rule-based Expert Systems and Machine Learning pipelines in Python (sklearn, keras, tensorflow, pytorch)
- Experience with Deep Learning and Data fusion approaches
- Strong Experience with exploratory data analysis with R, Orange Data Mining Tool and Weka
- Strong Experience with Explainable AI (XAI) packages (shap, lime)

Bioinformatics

- Experience with Next-Generation Sequencing data
- Experience with Single-Cell RNA sequencing data
- Experience with variant annotation pipeline (VEP, ANNOVAR)

Scientific skills

- Computer programming
- Expertise in data mining, artificial intelligence, machine learning and statistics
- Expertise in biomedical signals and images elaboration
- Scientific writing

Computer skills

General Expertise

Excellent Computer programming, Operative systems, bash programming, git

Good Databases, text elaboration, GUI implementation

Languages

Excellent PYTHON, MATLAB, R

Good C, JAVA, REACT JS,REACT HTML, CSS, PERL, SQL, MYSQL, MONGODB, LaTeX

Known Integrated Development Environment

Excellent PyCharm, RStudio

Good Spyder, Eclipse

Known scientific tools

Excellent Orange data mining, Anaconda Platform

Good Weka

Known general purpose tools

Microsoft Office, LibreOffice, GIMP, ImageJ

Presentations

2023 **Panelist** and **Oral presentation**: "Trustworthy AI for Healthcare: what does it mean and where are we at?", AMIA Informatics Summit, March 16th, Seattle (WA)

2022 **Oral presentation** at the MIE (Medical Informatics Europe) conference, Nice (France)

2022 **Poster presentation** at the ESHG (European Society of Human Genetics) conference, Vienna (Austria)

2021 **Oral presentation** at SMARTERCARE, a workshop of the Italian Association of Artificial Intelligence (AIxIA) virtual conference

2020 **Oral presentation** at AMIA (American Medical Informatics Association) Annual Symposium 2020 (virtual)

2019 **Oral presentations** at the 17th AIME (Artificial Intelligence In Medicine in Europe) Conference (Poznan, Poland)

2019 **Poster** at AMIA (American Medical Informatics Association) Informatics Summit, San Francisco (CA)

2018 **Oral presentation** at AMIA (American Medical Informatics Association) Informatics Summit, San Francisco (CA)

Selected research activities

2020-2023 **Development of reliability and explainability (XAI) approaches for trustworthy AI systems.** I am involved in the implementation of methods to assess the reliability of Machine Learning prediction, intended as the degree of trust that the ML prediction on a single case is correct. I am involved in the development of method for local XAI (named AraucanaXAI)

Project in collaboration with *Department of Medical Informatics, Amsterdam UMC, University of Amsterdam, the Netherlands* and *Laboratorio di Informatica Medica "Mario Stefanelli", University of Pavia*

2021-2023 **Phenotype-based genomic variant prioritization**. Variant prioritization tools can support Rare Disease patients diagnosis by pinpointing the causative variant(s) among thousands of genomic variations detected in a patients sample after sequencing. Within this project, I am responsible for the implementation of the ML system for variant prioritization. Our solution resulted in one of the best performing in the NIH-funded CAGI6 RGP challenge, where several teams, both from industry and academia, where asked to identify the causative variant in patients with Rare Diseases.

Project in collaboration with *enGenome*

2021-2023 **Early detection of Sars-cov2 variants through Machine Learning** I developed incremental learning and anomaly detection approaches to early detect new Sars-cov-2 variants from Spike protein sequences.

Project in collaboration with *University of Florida, US*, funded by NIH

2017–2020 **Methodologies and technologies of biomedical informatics and bioinformatics supporting the project "Genomic Profiling of rare hematological malignancies, development of personalized medicine strategies and their implementation into Rete Ematologica Lombarda (REL) clinical network** The aim of this project is to develop pipelines and tools for supporting clinical decision making in the hematological context.

Project funded by FRRB, Fondazione Regionale per la Ricerca Biomedica.

2019–2020 **Single-Cell Data analysis and Cell Fate Prediction** Statistical and Machine Learning algorithms have been applied to analyse Single-Cell gene expression data from Acute Myeloid Leukemia patients, to predict current and future state of the cell

Project in collaboration with *Bioinformatics Lab*, Ljubljana, Slovenia.

2018–2021 **Development of Machine Learning approaches for variant oncogenicity prediction**. I developed a pipeline for somatic variant oncogenicity prediction. Such pipeline takes as input VCF files, it annotates variant through the VEP annotator, and it uses annotation features to predict the oncogenicity of each variant based on Deep Learning and Machine Learning algorithms.

Teaching activity

2023, June 23rd Master Biologia e Biotecnologie della riproduzione: lessons "Sistemi di supporto per le decisioni cliniche" and "Ampliare la visione: bioingegneria e medicina della riproduzione"

Institution: *Università degli Studi di Pavia, Dipartimento di Medicina Interna e Terapia Medica*

2023, May 23rd Invited seminar, University of Florida: A primer on prediction model design, implementation and evaluation. Reference professor: Simone Marini

2021 Master Genetica Oncologica: lesson "Bioinformatics solutions to support Genomic Variant Interpretation in Precision Oncology". Reference professor: Riccardo Bellazzi

Institution: *Università degli Studi di Pavia, Dipartimento di Medicina Interna e Terapia Medica*

2020-2021 **Python Teaching Assistant** for the *Medicine Enhanced by Engineering Technologies* (MEET) course

Institution: *Università degli Studi di Pavia, Faculty of Medicine*

Reference professor: *Prof. Cristiana Larizza*

2018–2022 **Biomedical Informatics Tutor (Matlab)** for biomedical engineering courses

Institution: *Università degli Studi di Pavia, Faculty of Engineering*

Reference professor: *Prof. Lucia Sacchi, Prof. Riccardo Bellazzi*

2018–2021 Internet and Medicine Tutor for biomedical engineering courses

Institution: *Università degli Studi di Pavia, Faculty of Engineering*

Reference professor: *Prof. Giordano Lanzola*

2016- **Co-supervisor** of 3 bachelor and 4 master's theses in bioengineering

Institution: *Università degli Studi di Pavia*

Reference professors: *Prof. Riccardo Bellazzi*, *Prof. Paolo Magni*

Other activities

Reviewer for

- the Journal of Biomedical Informatics (JBI)
- Human Mutation
- The Journal of the American Medical Informatics Association (JAMIA)
- Bioinformatics Advances
- BioData Mining
- BMC Supplements
- Plos One
- GENE
- Microorganisms
- Conferences: Conference of Information and Knowledge Management 2020, AMIA Informatics Summits 2019 and 2020, AMIA Annual Symposium 2023

Program Committee Member of the Special Session Actionable Explainable AI (AxAI) inthe Cross Domain Conference for Machine Learning and Knowledge Extraction (CD-MAKE 2023)

Languages

Italian Mother tongue

English Professional working proficiency

International and national organization, Awards

2021 National prize for doctoral thesis in the Bioengineering area, National Bioengineering Group (GNB), Italy

2021– DAAD AInet fellow. This fellowship by the German Academic Exchange Service recognizes early-career researchers in the field of Artificial Intelligence

2021-2022 Member of the Associazione Italiana Intelligenza Artificiale (AIxIA)

2018–2019 Student Member of the American Medical Informatics Association (AMIA)

2019 Founder student fellow of the *National Bioengineering Group* (GNB) association, Italy

Relevant Coursework

2021 Deep Neural Networks with Pytorch (IBM online course provided by Coursera) (`https://coursera.org/share/482f88863504891feb2411f1b58abf33`)

2022 Introduction to Machine Learning in Production (DeepLearning.AI course provided by Coursera)(`https://coursera.org/share/6096c9beb994408e98f5c512e798b74f`)

Patent

2020 European Patent N. WO2022029567 : A METHOD FOR DETERMINING THE PATHOGENICITY/BENIGNITY OF A GENOMIC VARIANT IN CONNECTION WITH A GIVEN DISEASE (Italian Patent N. 102020000019180)

Selected Publications

Journals

2023 E. Parimbelli, T. M. Buonocore, - **G. Nicora**, W. Michalowski, S. Wilk, R. Bellazzi. Why did AI get this one wrong?—Tree-based explanations of machine learning model predictions. *Artificial Intelligence in Medicine*, https://doi.org/10.1016/j.artmed.2022.102471, Jan 2023

2022 - **G. Nicora**, S. Zucca, I. Limongelli, R. Bellazzi, P. Magni. A machine learning approach based on ACMG/AMP guidelines for genomic variant classification and prioritization *Scientific Reports*, https://doi.org/10.1038/s41598-022-06547-3, Feb 2022

2022 - **G. Nicora**, M. Rios, A. Abu-Hanna, R. Bellazzi. Evaluating pointwise reliability of machine learning prediction *Journal of Biomedical Informatics*, https://doi.org/10.1016/j.jbi.2022.103996, Mar 2022

2020 - **G. Nicora**, F. Vitali, A. Dagliati, N. Geifman, R. Bellazzi. Integrated Multi-Omics Analyses in Oncology: A Review of Machine Learning Methods and Tools. *Frontiers in Oncology*, https://doi.org/10.3389/fonc.2020.01030, June 2020

- **G. Nicora**, F. Moretti, E. Sauta, M. Della Porta, L. Malcovati, M. Cazzola, S. Quaglini, R. Bellazzi, "A continuous-time Markov model approach for modeling myelodysplastic syndromes progression from cross-sectional data", *Journal of Biomedical Informatics*, Volume 104, April 2020, 103398, https://doi.org/10.1016/j.jbi.2020.103398

- G. Coticchio, G. Fiorentino, **G. Nicora**,R. Sciajno,F. Cavalera R. Bellazzi,S. Garagna, A. Borini, M. Zuccotti, Harnessing cytoplasmic particles movement of the human early embryo analysed by advanced imaging and artificial intelligence to predict development to blastocyst stage, Reproductive BioMedicine Online, December 2020, https://doi.org/10.1016/j.rbmo.2020.12.008

2018 - **G. Nicora**, Ivan Limongelli, Patrick Gambelli, Mirella Memmi, Carlo Napolitano, Alberto Malovini, Andrea Mazzanti, Silvia Priori, Riccardo Bellazzi, "CardioVAI: An automatic implementation of ACMG-AMP variant interpretation guidelines in the diagnosis of cardiovascular diseases", *Human Mutation*, October 2018, PMID:30298955, https://doi.org/10.1002/humu.23665

Conference Proceedings

2021 - **G. Nicora**, R. Bellazzi. A Reliable Machine Learning Approach applied to Single-Cell Classification in Acute Myeloid Leukemia. AMIA Annu Symp Proc. 2021 Jan 25;2020:925-932. PMID: 33936468; PMCID: PMC8075526.

2019 - **G. Nicora**, Simone Marini, Ivan Limongelli, Ettore Rizzo, Stefano Montoli, Francesca Tricomi, Riccardo Bellazzi, A Semi-supervised Learning Approach for Pan-Cancer Somatic Genomic Variant Classification. In: Riaño D., Wilk S., ten Teije A. (eds) Artificial Intelligence in Medicine. AIME 2019. Lecture Notes in Computer Science, vol 11526. Springer, Cham

- **G. Nicora**, Ivan Limongelli, Riccardo Cova, Matteo Giovanni Della Porta, Luca Malcovati, Mario Cazzola, Riccardo Bellazzi, A Rule-Based Expert System for Automatic Implementation of Somatic Variant Clinical Interpretation Guidelines, In: Riaño D., Wilk S., ten Teije A. (eds) Artificial Intelligence in Medicine. AIME 2019. Lecture Notes in Computer Science, vol 11526. Springer, Cham

- G. Coticchio, R. Sciajno, G. Fiorentino, F. Cavalera **G. Nicora**, R. Bellazzi, A. Borini, S. Garagna, M. Zuccotti, Artificial neural-network analysis combined with time-lapse imaging predicts embryo ability to develop to the blastocyst stage, Fertility and Sterility, Volume 112, Issue 3, Supplement, September 2019, Pages e273-e274 https://doi.org/10.1016/j.fertnstert.2019.07.810

Conference Abstracts

2022 **G. Nicora**, F. De Paoli, I. Limongelli, P. Magni, Performance of ACMG/AMP guidelines to tackle digenic variant interpretation: do we need a digenic variant interpreter?, Europen Society of Human Genetics (ESHG) Annual Meeting, June 2022

2020 - F. De Paoli, I. Limongelli, E. Rizzo, **G. Nicora**, P. Magni, An automatic implementation of ACMG/ClinGen guidelines for constitutional Copy Number Variants annotation and interpretation, American Society of Human Genetics (ASHG) Annual Meeting, October 2020

2019 - **G. Nicora**, F. Moretti, E- Sauta, L. Malcovati, M. Della Porta, S. Quaglini, M. Cazzola, R- Bellazzi, "A countinuous-time Markov approach for modelling myelodysplastic syndromes progression from cross-sectional data", poster presentation at AMIA (American Medical Informatics Association) 2019

- **G.Nicora**, I. Limongelli, S. Zucca, R. Santolisier, P. Magni, R. Bellazzi, A comparison of eVAI, CADD and VVP variant prediction results on the ICR639 hereditary cancer dataset, abstract for poster presentation, American Society of Human Genetics (ASHG) Annual Meeting, Houston (US-TX), october 2019

- **G. Nicora**, Ivan Limongelli, Patrick Gambelli, Mirella Memmi, Carlo Napolitano, Alberto Malovini, Andrea Mazzanti, Silvia Priori, Riccardo Bellazzi, "An automatic implementation of ACMG/AMP variant interpretation guidelines.", abstract ESHG (European Society of Human Genetics) June 2018.

- **G. Nicora**, I. Limongelli, P. Gambelli, M. Memmi, C. Napolitano, A. Malovini, A. Mazzanti, S. Priori, R. Bellazzi, "A Rule-based Expert System for automatic genomic variant interpretation", abstract and poster presentation GNB (Gruppo Nazionale Bioingegneria), Milan (Italy) June 2018. (Poster Presentation)

2017 - I. Limongelli, **G. Nicora**, P. Gambelli, M. Memmi, C. Napolitano, A. Malovini, A. Mazzanti, S. Priori, R. Bellazzi, "An automated guidelines-based approach for variants pathogenicity assessment in the diagnosis of genetic cardiovascular diseases.", abstract for poster presentation SIGU (Società Italiana Genetica Umana), Naples (Italy) November 2017. (Poster presentation)

Pavia, August 28, 2023