6B Radcliffe Road
NG2 5FW Nottingham
United Kingdom
07749230137
jeremieclos@gmail.com
jclos.github.io

Jérémie Clos

Education

2012–2019 **PhD**, *Robert Gordon University*, Aberdeen.
PhD in Artificial Intelligence

2011–2012 **MSc**, *University of Toulouse*, Toulouse, *With Distinction.*
MSc in Information Retrieval and Artificial Intelligence

2010–2011 **Hons**, *Robert Gordon University*, Aberdeen, *2:1.*
Bachelor (Hons) in Computing

2006–2010 **BSc**, *University of Toulouse*, Toulouse.
Bachelor in Computer Science

PhD research

Title *Representation and Learning Schemes for Argumentation Mining*

Supervisors Prof. Nirmalie Wiratunga, Dr. Stewart Massie, Prof. Joemon Jose, Dr. Guillaume Cabanac

Key topics machine learning, opinion mining, stance detection

Description My PhD project focuses on the study of representation and learning schemes for argumentation stance mining from social media content, in order to enable fine-grained debate analysis. Funding for this research was provided by a SICSA prize studentship in the themes of Next Generation Internet and Modelling and Abstraction. I particularly focus on the use of a shallow computational graph for stance lexicon extraction.

Master thesis

Title *Predicting Entry Points in Q&A Systems*

Supervisors Dr. Guillaume Cabanac, Prof. Mohand Boughanem, Dr. Henri Prade

Key topics machine learning, information retrieval, social media mining

Description This work is an investigation in engineering features to enhance entry point prediction in Q&A systems. We focus on modelling answer complexity using compression ratios. This research was done in the information retrieval team of the University of Toulouse computer science research laboratory.

Research experience

07/2018–current **Research associate**, *University of Nottingham*, Nottingham.
Research associate role on the ENLIVEN project. Responsibilities include:
- Developing the ENLIVEN IDSS prototype.
- Disseminating research about case-based reasoning for decision support systems.
- Investigating state-of-the-art techniques for knowledge representation in a budgeted learning context.

09/2017–12/2017 **Research assistant**, *Robert Gordon University*, Aberdeen.

My work was focused on a DataLab-funded project with the National Engineering Laboratory (NEL). My main tasks were:

- Performing a systematic study of data analysis techniques for drift prediction in flow modelling.
- Investigating the robustness of modern machine learning methods for extremely budgeted learning, with training sets restricted to a dozen data points.
- Developing a R toolkit to facilitate the exploration and extension of the current methodology of NEL. We presented the toolkit during our closing meeting in December 2017.

09/2015–04/2017 **Data scientist/applied researcher**, *Cognitive Geology*, Edinburgh.

As the core research scientist in a small team composed of petroleum geologists and software engineers, I was in charge of the research and development for our product. Most of my work focused on the practical applications of generalized additive models for expert-driven geological modelling.

- Investigation of modern regression techniques for geological modelling.
- Research and implementation of heuristic-constrained logistic fitting algorithms to analyse geological data.
- Research and implementation of fast grid splitting algorithms for geological modelling.

Working both at the prototyping and implementation phases, I used the following technologies:

- R/Caret and Python/Scikit-learn for prototyping algorithms.
- C#, Syncfusion and Accord.NET for final implementation.

More stuff about the project goes here.

03/2009–06/2012 **Research assistant**, *IRIT Research Lab*, Toulouse (France).

As an undergraduate research assistant, I participated in multiple research projects both on a national and European scale, involving university and industrial partners.

- **Summer 2012**: Multiple small projects for members of the information systems team: paper translation, literature review, and prototype development.
- **Summer 2011**: METHODEO project, on the evaluation of distributed indexing systems for large multimedia archives for large-scale surveillance.
- **Summer 2010**: LINDO project, on the implementation of an open system for indexing and retrieval of multimedia objects in large distributed archives.
- **Spring–Summer 2009**: CERISE project, on the implementation of a systems to systems interface engineering, and study of impact analysis and data traceability in large-scale engineering projects.

Teaching experience

07/2018–current **Part-time teaching assistant**, *University of Nottingham*, Nottingham.

- **Fundamentals of Artificial Intelligence** (level 1). An introductory first year module regarding AI, its history, and its key techniques.
- **Programming Paradigms** (level 1), focused on teaching the basic principles of program design and implementation using the object-oriented and functional approaches to programming.
- **Data Modelling and Analysis** (level 4). An advanced module that focuses on the principles, techniques and applications of a range of data analysis and modelling techniques.

Additional teaching duties:

- Bachelor thesis supervision

01/2018–07/2018 **Full-time teaching assistant**, *University of Nottingham*, Nottingham.

As full-time teaching staff, I intervened on several undergraduate and postgraduate modules. My role was focused on running practical labs, helping students in guided tutorials and marking exam papers, which involved a large part of face to face teaching. I was assigned to the following modules:

- **Databases and Interfaces** (level 1). An introductory module to the theory and practice of modern database systems.
- **Fundamentals of Artificial Intelligence** (level 1). An introductory first year module regarding AI, its history, and its key techniques.
- **Programming Paradigms** (level 1), focused on teaching the basic principles of program design and implementation using the object-oriented and functional approaches to programming.
- **Software Engineering** (level 1), aimed at providing a general understanding of software engineering, and the typical phases of the software lifecycle.
- **Artificial Intelligence Methods** (level 2), building on the first year AI module in order to further an appreciation of various AI techniques.
- **Computer Graphics** (level 3), focused on providing students with practical means for learning and understanding the 3D computer graphics pipeline, methods and programming skills for modelling, transforming, viewing and projecting 3D objects, and adding realism to them.
- **Computer Security** (level 3), focusing on security issues associated with computers and computer networks.
- **Data Modelling and Analysis** (level 4). An advanced module that focuses on the principles, techniques and applications of a range of data analysis and modelling techniques.
- **Games** (level 4). An advanced module teaching the history of games as an industry and a form of entertainment; the principles of game design and implementation, and the design and development of games.
- **Software Engineering Management**, (level 4) focusing on preparing students for professional software development, software project management and software and IT project specification and development.

Additional teaching duties:
- Group project supervision
- Delivery of tutorials
- Masters thesis supervision

09/2012–06/2014 **Teaching assistant**, *Robert Gordon University*, Aberdeen.

As a graduate teaching assistant, I intervened on a series of undergraduate and postgraduate modules. My role was focused on running practical labs, helping students in guided tutorials and marking exam papers, which involved a large part of face to face teaching.

- **MSc modules:**
  - **Adaptive Intelligent Systems**, focusing on applied optimization using nature-inspired metaheuristics (genetic algorithms, particle swarm optimisation).
  - **Data Mining**, focusing on the use of standard data mining algorithms to discover knowledge from datasets, focused on business applications such as churn prediction.

- **Undergraduate modules:**
  - **Database Systems** (3rd and 4th year students), a module on SQL, Oracle database systems administration and PL/SQL programming, focused on applied skills in database design.
  - **Introduction to Databases** (2nd and 3rd year students), a module on SQL and MSSQL Server, focused on introducing students to database design and manipulation.

Supervisory experience

2018–2019 Bachelor thesis: A comparative study of machine learning algorithms on mental workload estimation.

2018–2019 Bachelor thesis: An online bill splitting system with neural network projection.

2017–2018 MSc thesis: An affordable system for sleep apnoea monitoring.

Administrative experience

2017–2018 Main organiser of the RGU-AI research weekly team seminars, administrator of the associated mailing list and Slack channel.

2017 Member of the organising and bidding team of the 2017 European Conference on Information Retrieval (ECIR).

2015 Member of the organising team of the 2015 SICSA workshop on Mining Social Media.

2014–2015 Organiser of the monthly IDEAS PhD Computer Science student seminars.

Presentations

Paper presentations
IJCAI-XAI 2017 - IJCAI Workshop on Explainable AI.
TPDL 2017 - Theory and Practice of Digital Libraries.
ECIR 2017 - European Conference in Information Retrieval.
LDK 2017 - Language, Data and Knowledge.
ECA 2015 - European Conference on Argumentation.
SWAM 2014 - SICSA Workshop on Argument Mining.

Poster presentations
ECIR 2017 - European Conference in Information Retrieval.
SICSA PhD conference 2015.
IDEAS symposium 2015.
SICSA PhD conference 2014.
IDEAS symposium 2014.
SICSA PhD conference 2013.

Tutorials RGU-AI research seminars: "Introduction to neural networks".

Miscellaneous IDEAS PhD student seminar.

Awards

2014 SICSA Summer School bursary: funding to attend the SICSA Summer School on Argumentation.

2014 SICSA Travel funding: travel and accommodation funding for the Scottish Workshop on Argument Mining.

2013 SICSA Travel grant: travel funding for the Scottish Workshop on Information Retrieval.

2015 SICSA Prize Studentship: 3-year funding for PhD in a SICSA university.

2015 EU scholarship: European funding to attend the Summer School in Information Foraging

Programming skills

Python Machine learning research and development. Experience with Scikit-learn, Numpy and Tensorflow for algorithm development.

Java	Teaching, research and general development.
C#	General development using the Syncfusion graphical package and the Accord.NET machine learning library.
R	Statistical analysis and generating (beautiful) plots.
SQL	Mostly teaching (undergraduate students in year 2 and 3).

General computing skills

Version control	Git, TFS
Databases	Oracle, MSSQL
Python toolkits	Scikit-learn, Tensorflow
R toolkits	Caret, ggplot2
Scientific writing	LaTeX, Markdown
Others	Unity (teaching), Unreal engine (teaching)
Markup languages	HTML, XML

Scientific interests

Machine learning	Neural networks, recurrent networks, attention models, meta-learning.
Social media mining	Social media argumentation, debate mapping, argument mining.
Digital health	Monitoring and management of mental health using natural language processing.
Information retrieval & management	Personal information management, scientometrics.

Languages

French	Native.
English	Fluent.
Latin	Intermediate.
Italian	Beginner.

References

(References are listed by decreasing level of recency of employment.)

[1] Prof. John McCall
Head of the RGU School of Computing Science and Digital Media.
j.mccall@rgu.ac.uk
+44 (0) 1224 262780

[2] Prof. Nirmalie Wiratunga
Main PhD advisor.
n.wiratunga@rgu.ac.uk
+44 (0) 1224 262573

[3] Dr. Guillaume Cabanac
Main MSc advisor, external PhD advisor.
guillaume.cabanac@univ-tlse3.fr
+33 (0) 5615 57273

[4] Prof. Mohand Boughanem
MSc advisor.
mohand.boughanem@irit.fr
+33 (0) 5615 57416

Publications

[1] Jérémie Clos, Anil Bandhakavi, Nirmalie Wiratunga, and Guillaume Cabanac. Predicting emotional reaction in social networks. In *European Conference on Information Retrieva*l, pages 527–533. Springer, Cham, 2017.

[2] Jérémie Clos and Nirmalie Wiratunga. Lexicon induction for interpretable text classification. In *International Conference on Theory and Practice of Digital Libraries*, pages 498–510. Springer, Cham, 2017.

[3] Jérémie Clos and Nirmalie Wiratunga. Neural induction of a lexicon for fast and interpretable stance classification. In *International Conference on Language, Data and Knowledge*, pages 181–193. Springer, Cham, 2017.

[4] Jérémie Clos, Nirmalie Wiratunga, Joemon Jose, Stewart Massie, and Guillaume Cabanac. Towards argumentative opinion mining in online discussions. In *Proceedings of the SICSA Workshop on Argument Mining*, page 10, 2014.

[5] Jérémie Clos, Nirmalie Wiratunga, and Stewart Massie. Towards explainable text classification by jointly learning lexicon and modifier terms. In *IJCAI Workshop on Explainable Artificial Intelligence*, 2017.

[6] Jérémie Clos, Nirmalie Wiratunga, Stewart Massie, and Guillaume Cabanac. Shallow techniques for argument mining. In *European Conference on Argumentation* (ECA 2015), volume 63, pages pp–341, 2015.

[7] Kyle Martin, Nirmalie Wiratunga, Sadiq Sani, Stewart Massie, and Jérémie Clos. A convolutional siamese network for developing similarity knowledge in the SelfBACK dataset. 2017.

[8] Kyle Martin, Nirmalie Wiratunga, Sadiq Sani, Stewart Massie, and Jérémie Clos. Informed pair selection for self-paced metric learning in Siamese neural networks. 2018.

Technical reports

[TR1] Jérémie Clos, André Péninou, and Florence Sèdes. Projet Cerise - Rapport de synthèse : Etat de l'art sur l'analyse d'impact et la traçabilité. Rapport de contratIRIT/RR–2009-23–FR, IRIT, Université Paul Sabatier, Toulouse, June 2009.

[TR2] André Péninou, Florence Sèdes, and Jérémie Clos. Projet Cerise - Rapport de synthèse : Synthèse pour l'élaboration d'un référentiel d'ingénierie des interfaces entre systèmes. Rapport de contrat IRIT/RR–2009-24–FR, IRIT, Université Paul Sabatier, Toulouse, June 2009.

[TR3] John McCall, Olivier Regnier-Coudert, and Jérémie Clos. Predictive modelling of multivariate flow measurements for analysis of meter condition. Technical report, Robert Gordon University, Robert Gordon University, Aberdeen, December 2017.