CHEMSEDDINE BERBAGUE

Data Scientist | Expert Recommender Systems

8.chems@gmail.com +213-697-082836 Annaba, Algeria
8-chems 8-chems bchems Chemseddine_Berbague
https://8-chems.blogspot.com

PROFESSIONAL SUMMARY

Data-driven PhD in Computer Science with 9 years of experience in academia and industry, specializing in recommender systems, machine learning, large language models (LLMs), and semantic data modeling - key trends shaping modern data science. Proficient in Python, Java, SQL, and frameworks like Scikit-learn, PyTorch, Apache Jena, with expertise in OWL, RDFS, SPARQL for knowledge graphs. Authored/contributed 6 peer-reviewed publications with 500+ reads on ResearchGate and 50+ citations on Google Scholar. Shared insights via Kaggle notebooks and Blogspot posts. Experienced in teaching data science, mentoring innovative projects, and delivering scalable solutions for industry challenges.

KEY SKILLS

- **Programming**: Python (Pandas, NumPy, Scikit-learn), Java, MATLAB, C#, R
- **Data Modeling & Knowledge Graphs**: OWL, RDFS, SPARQL, Apache Jena
- **Data Science Tools**: Weka, Mahout, RankSys, Kaggle, LLMs, PyTorch, TensorFlow, Matplotlib.
- **Machine Learning**: Recommender Systems, Clustering, Deep Learning, Evolutionary Algorithms
- **Data Visualization**: Matplotlib, Visualization Tools
- **Database Management**: MySql, PostgresSQL, Access.
- **Web Development**: Vue.js, React, Redux
- **Project Management**: Jira, Confluence, Scrum/Kanban, Git
- **Soft Skills**: Academic mentoring, technical blogging, stakeholder communication, Agile methodologies

PROFESSIONAL EXPERIENCE

Higher School of Computer Science and Technologies (ESTIN)
Béjaïa, Algeria

Assistant Professor
Sep 2022 – Present

- Delivered **Foundations of Data Science** course to 60+ students annually, covering data acquisition, preprocessing, modeling, visualization, and communication using Python and Matplotlib.
- Taught **Information System Security Audit** course, training 50+ students in security fundamentals and audit methodologies.
- Supervised master's projects on LLM-based recommender systems and automated data science tools.
- Mentored PhD thesis on psychological impacts of recommender system manipulations, developing simulation scenarios and mitigation strategies.

Nechma Technology
Annaba, Algeria

Data Management Engineer
Oct 2020 – Mar 2022

- Developed an accounting module for the Algerian Accounting System using Python and XML, streamlining financial processes.
- Designed configurable reporting solutions for tax compliance.
- Automated data verification with checklists, enhancing data accuracy.
- Optimized user interfaces.

Self-Employed
Annaba, Algeria

Freelance Developer
Oct 2019 – Sep 2020

- Built dynamic web and desktop applications using Vue.js, React, Redux, and SQL databases.
- Developed Java-based backend solutions for efficient data storage and retrieval.

Badji Mokhtar University
Annaba, Algeria

Part-Time Assistant Professor
Jan 2016 – Jun 2021

- Mentored 50+ students in MATLAB programming labs, enhancing their skills in data structures, matrix calculations, and visualization.

- Guided students from basic to advanced programming concepts, improving project completion rates by 30%.

RESEARCH INTERNSHIPS

University of Bari
Bari, Italy

Research Intern
Dec 2024

- Developed an AI-based platform using multimodal LLMs for environmental analysis, integrating geospatial data and RAG for non-expert user insights.
- Designed a conversational interface for soil and water management analytics.

University of Bolzano
Bolzano, Italy

Research Intern
Dec 2019 – Dec 2020

- Contributed to session-based and multi-objective recommender systems, resulting in 2 publications.
- Developed genetic algorithms for optimizing recommendation accuracy, diversity, and novelty.

EDUCATION

Badji Mokhtar University
Annaba, Algeria

Ph.D. in Computer Science
Jan 2016 – Jun 2021

- Thesis: "Study of Novelty and Diversity in Collaborative Filtering Recommendation Approaches"
- Developed an evolutionary optimization algorithm balancing diversity, novelty, and relevance in recommendations.

Badji Mokhtar University
Annaba, Algeria

Master's in Information Systems
Sep 2013 – Jun 2015

- Final Project: "Leveraging Semantic Web with Association Rules to Enrich Information Systems"
- Built a Java-based visual tool for querying OWL/RDFS ontologies using SPARQL and association rule mining.

PROJECTS

Kaggle Community Contribution
Remote

Kaggle Recommender System Notebook
2023

- Published various public Kaggle notebooks demonstrating data preprocessing, exploratory data analysis, model building, model validation, results communication, and storytelling for different machine learning applications such as clustering and classification. recommender system, achieving 10+ community upvotes.
- Focused on statistical based techniques such as feature engineering, dimensionality reduction, data smoothing, data cleansing, etc.

Blogspot Contribution
Remote

Data Science Blog
2024 – Present

- Authored posts in various IT domains, including machine learning, experimental design, LLMs, recommender systems, data visualization, and more.
- Explained complex concepts like clustering and model evaluation for a general audience, enhancing community knowledge sharing.

Master's Project
Annaba, Algeria

Semantic Web Knowledge Extraction Tool
2015

- Developed a visual tool for exploring and querying OWL/RDFS ontologies, (i.e, using Java, Apache Jena, and SPARQL), integrating association rule mining to extract semantic knowledge.
- Enabled efficient filtering and visualization of complex datasets, enhancing data accessibility for users.

Research Project

Annaba, Algeria

Evolutionary Recommender System

2022

- Developed an evolutionary algorithm for recommender systems, balancing accuracy, diversity, and novelty, published in *Int. J. Bus. Intell. Data Min.* (2022).
- Achieved 2% (i.e., on average) improvement in recommendation diversity on a 1M-movielens dataset.

Engineering Projects Supervison

Béjaïa, Algeria

ESTIN

2022 – 2025

- **LLM-Based Data Science Tool:** Supervised the development of an automated end-to-end tool using LLMs for data preprocessing, visualization, modeling, and validation.
- **LLM-Based Recommender System:** Guided the creation of a personalized recommendation system integrating LLMs for news recommendation.
- **Health Data Forcasting Using Diffusion Models:** Mentored a project building a classifier that predicts health status.
- **Visual-Based User Behavior Modeling System for Recommendation:** Supervised a project exploring visual data and deep contrastive clustering to provide accurate recommendations.

Foundations of Data Science Course

Béjaïa, Algeria

The course include various theoritical and practical aspects related to data science:

2023

- Introduction to data science methodologies, and presenting its different stages, tools, and requirements.
- Presenting different data processing, cleansing methods, and feature engineering techniques.
- Enumerating various statistical-based methods to summarize, explain, and prune data.
- Citing data science applications, examples, tools, and applying practical activities.
- Providing mentorship for data communication, visualization, and interpretation.

PUBLICATIONS

- **Berbague, C.**, Seridi-Bouchelaghem, H., Karabadji, N. E. I., Symeonidis, P., Zanker, M., "An evolutionary-based approach for providing accurate and novel recommendations," *Int. J. Bus. Intell. Data Min.*, vol. 21, no. 2, pp. 129–148, 2022.
- Symeonidis, P., Chaltsev, D., **Berbague, C.**, Zanker, M., "Sequence-aware news recommendations by combining intra- with inter-session user information," *Inf. Retr. J.*, vol. 25, no. 4, pp. 461–480, 2022.
- **Berbague, C.**, Karabadji, N. E. I., Seridi, H., Symeonidis, P., Manolopoulos, Y., Dhifli, W., "An overlapping clustering approach for precision, diversity and novelty-aware recommendations," *Expert Syst. Appl.*, vol. 177, p. 114917, 2021.
- Symeonidis, P., Bellinazzi, L., **Berbague, C.**, Zanker, M., "Safe and effective recommendation of drug combinations based on matrix co-factorization," in *2023 IEEE 36th Int. Symp. on Computer-Based Medical Systems (CBMS)*, 2023, pp. 634–639. DOI:10.1109/CBMS58004.2023.00292
- **Berbague, C.**, Karabadji, N. E. I., Seridi, H., "An evolutionary scheme for improving recommender system using clustering," in *CIIA 2018*, 2018, pp. 290–301.
- **Berbague, C.**, Karabadji, N. E. I., Seridi, H., "Enhancing the sales diversity using a two-stage improved KNN algorithm," in *MISC 2018*, 2018, pp. 193–203.
- **Berbague, C.**, Karabadji, N. E. I., Seridi, H., "Recommendation diversification using a weighted similarity measure in user-based collaborative filtering," in *ISPS 2018*, 2018, pp. 1–6. DOI:10.1109/ISPS.2018.8379011

PROFESSIONAL DEVELOPMENT

- Conference Reviewer: RecSys (2020), ECAI (2020), TheWebConf (2020), ACM UMAP (2019).
- Journal Reviewer: *Information Systems*.

REFERENCES

Prof. Panagiotis Symeonidis
symeon@aegean.gr
University of the Aegean, Greece

Prof. Youssef Elmir
elmir@estin.dz
ESTIN, Béjaïa, Algeria

Dr. Harriri Walid
walid21548@hotmail.fr
UBMA, Annaba, Algeria