AMIN FARVARDIN

Desired post: *R&D position in Data Science | Data Engineer*
Young Doctor benefit from the CIR (research tax credit)

@ m.farvardin@hotmail.com   📞 +33 (0)7 69 11 12 20   📍 Courbevoie, France

in linkedin.com/in/FarvardinA

🔗 lamsade.dauphine.fr/∼afarvardin   K https://www.kaggle.com/aminfar

🔗 https://github.com/afarvardin   🔗 https://git.lamsade.fr/afarvardin

ACTUAL SITUATION

Big Data Architect Certification - Online Courses (Instructor)

Université Paris-Dauphine | Tunis

📅 June 2021 – ongoing
📍 Paris, France

– Data Processing
  — Data lake / LakeHouse / Delta
– Stream Processing
  — Lambda and Kappa Architectures

Lecture and Research Assistant

LAMSADE, Université Paris-Dauphine

📅 Sep 2019 – ongoing
📍 Paris, France

– Systemes de Langages de Programmation pour le "Big Data"
– Large Scale Machine Learning in Apache Spark
– Mise à niveau en Java
– Database (RDBMS)

EXPERTISE AND INTERESTS

Data Architecture
– Batch Processing, Stream Processing, Micro-Batch Processing
– Lambda and Kappa Architecture

Statistical Learning Basis
– Supervised Method: Linear regression, Logistic regression, SVM, KNN, PCA, Random Forest, Naive Bayes
– Unsupervised Method: k-means, Hierarchical Clustering, LDA
– Deep Learning: ANN, CNN

WORK EXPERIENCE

Data Science Pole in LAMSADE (Université-Paris Dauphine)

Data Streaming, Data Processing, Architecture Big Data

📅 Nov 2016 - Dec 2020
📍 Paris, France

Scalable Process of Streaming Massive RDF Data
– Devised, designed, and implemented a novel indexing technique in Spark Streaming for RDF saturation
– The proposed method is capable of identifying the relation of already past data with the current one
– Saturate Streaming RDF Incrementally
– Memory usage reduction by ∼40%
– Improve execution time by ∼17x
– Tuning the Spark Process and Partitions

`Spark & Spark Streaming` `Kafka` `Scala` `Java` `Python (Pandas, NumPy)` `Hadoop Yarn` `HDFS` `Azure Databricks` `AWS EMR/S3` `Google BigQuery` `Apache Airflow`

Full Stack Engineer

Text Mining

📅 Oct 2015 – Oct 2016
📍 Lille, France

Identification and analysis of land studies in TERRE-ISTEX project.
– Identification and analysis of land studies
– Annotate spatial text geographically

- Querying annotated spatial text

Text mining (GATE & GATE Embedded) · NLP · RESTful API · Java · Spring MVC · Elasticsearch · SQL · NoSQL · JS, HTML, CSS

Full Stack Engineer
Text Mining (Internship)

📅 April – Sep 2015
📍 Montpellier, France

Semantic annotation associated with land development projects

Text mining (GATE & GATE Embedded) · NLP · RESTful API · Java · Spring MVC · SQL · JS, HTML, CSS

Data Visualization
SIG L-R (Volunteer)

📅 Feb - April 2015
📍 Montpellier, France

Utilizing the registered participants' data to show what we can do with even none relevant data and how much knowledge potentially hidden in data

D3.js · JSON · JS

Full Stack Engineer
Text Mining (Master Internship)

📅 Jan 2014 – Jan 2015
📍 Montpellier, France

Semantic web services for opinion mining from land-use planning documents
- Annotate text by creating NLP rules
- Prepare a standalone pipeline
- Transform the standalone pipeline into a website application by utilizing GATE Embedded
- Implement a web-tool for analyzing land-use planning documents

Text mining (GATE & GATE Embedded) · NLP · RESTful API · Java · ASP.NET MVC - C# · Spring MVC · SQL · JS, HTML, CSS

Web Development on Azure Cloud
Building a RESTful web service

📅 Feb 2012 – July 2012
📍 Hyderabad, India

Design and implement a Phone-Book as a web service on Azure

ASP.NET · SQL · SOAP · Azure

INDUSTRIAL EXPERIENCE

Software Engineer
Petroleum Industry

📅 Sep 2006 – July 2008
📍 Shiraz, Iran

Design, implement, and maintain the software (logistics, repairs) for Iran's petroleum industry.

ASP.NET · C# · SQL

Back-end Web Developer
ESG Co. Ltd.

📅 June 2004 – June 2006
📍 Shiraz, Iran

ASP classic · ASP.NET · HTML, CSS, JS · CMS (dnn)

EDUCATION

Ph.D. in Streaming Big Data - LAMSADE (Université-Paris Dauphine)
Title: **Scalable Saturation of Streaming RDF Data**

📅 Nov'16 – Jan, 2021
📍 Paris, France

We present the first solution for reasoning over large streams of RDF data using big data platforms. In doing so, we focus on the saturation operation, which seeks to infer implicit RDF triples given RDF schema or OWL-Horst constraints. Indeed, unlike existing solutions which saturate RDF data in bulk, our solution carefully identifies the fragment of the existing (and already saturated) RDF dataset that needs to be considered given the fresh RDF statements delivered by the stream. Thereby, it performs the saturation in an incremental manner.

Master (Information Technology, JNTU)
Project: **Semantic web services for opinion mining from land-use planning documents**

Grade: First Class with Distinction

🗓 Sep 2012 - Oct 2014
📍 Hyderabad, India

Bachelor (Computer Engineering, JNTU)
Project: **A Phone-Book as a web service on Azure**

🗓 Sep 2008 - Jun 2012
📍 Hyderabad, India

Associate's Degree
Computer Software Engineering

🗓 2001 - 2004
📍 Shiraz, Iran

STRENGTHS

- **D**evelopment Languages, **F**rameworks, & **DB**

D: Scala | JAVA | Python (Pandas, NumPy) | C# & VB.NET | C

F: Spark | Hahoop Yarn | Kafka | Elasticsearch | AWS EC2, EMR, S3 & Lambda | Airflow | Docker | Kubernetes | ASP.NET MVC | Spring MVC

DB: NoSQL (HDFS, MangoDB, AWS S3 & DynamoDB, etc.) | MySQL, MSSQL | RDF/XML

- Skills

Stream & Batch Processing (Big Data) | Scrum | Text mining (NLP) | GATE(GUI, Embedded) | NLP | Git | Linux | Spring MVC | ETL concepts | Design Patterns | RESTful & SOAP | d3.js | LaTeX | JS, jQuery | HTML5, CSS3 | IIS, Glassfish 4.0

SOFT SKILLS

Adaptability | Flexibility | Responsibility | Team Work | Time management | Work Ethic | Leadership | Problem Solving | Self-Motivation | Listening | Personal Development

PUBLICATIONS (TOP 5)

1. Farvardin MA., Colazzo D., Belhajjame K., et al. "Scalable Saturation of Streaming RDF Triples", TLDKS 2020 (Journal).
2. Farvardin MA., Colazzo D., Belhajjame K., et al. "Streaming saturation for large RDF graphs with dynamic schema information", DBPL 2019, Phoenix, AZ, USA.
3. Kergosien E., Farvardin MA., et al. "Automatic Identification of Research Fields in Scientific Papers", LREC 2018, Japan.
4. Farvardin MA., Kergosien E., Roche M., Teisseire M. "A Webtool for Analyzing Land-use Planning Documents.", ISWC (Posters & Demos) 2015.
5. Farvardin MA., Kergosien E. "Senterritoire for Spatio-Temporal Representation of Opinions", NLDB 2014.

HONERS & AWARDS

Full PhD Scholarship at Université-Paris Dauphine
Minister of High Education of France

📅 2016
📍 Paris, France

Master II Scholarship in CIRAD
Minister of High Education of France

📅 2014
📍 Montpellier, France

MCSD
Web Application (Credential ID: E326-9247)

📅 Jan 2013
📍 Hyderabad, India

LANGUAGES

English
French
Persian

●●●●●
●●○○○
●●●●●