

# Batoul Sharafeddin

*Computational Scientist and Statistician*

## PERSONAL DETAILS

---

*Address* Bir Hassan Beirut, Lebanon  
*Phone* (+961) 78 863 528  
*Mail* bms19@aub.edu.lb

## EDUCATION

---

<b>MS Computational Science</b> <i>American University of Beirut</i> Cumulative GPA: 85	2017-2020
<b>BS Statistics</b> <i>American University of Beirut</i> Cumulative GPA: 81	2014-2017
<b>Minor Biology</b> <i>American University of Beirut</i>	2015-2017
<b>Lebanese Baccalaureate II in Life Sciences</b> <i>Choueifat International School</i>	2014

## WORK EXPERIENCE

---

<b>Research Assistant</b>	2017-2020
<ul style="list-style-type: none"><li>• <i>Pre-processing data</i></li><li>• <i>Working with different Natural Language Processing techniques</i></li><li>• <i>Classifying relevant words through a cross-document methodology</i></li><li>• <i>Implementing K-Means Clustering using R</i></li><li>• <i>Implementing Bayesian Graphs using R and Java</i></li><li>• <i>Implementing Hidden Markov Models using MatLab</i></li><li>• <i>Implementing Neural Networks using R and C</i></li><li>• <i>Implementing Distributional Similarity methods using Java</i></li><li>• <i>Developed a website using HTML, JavaScript, Python, and Java to display my work</i></li><li>• <i>Developing interactive methods in the website to allow user to update my results</i></li><li>• <i>Developing interactive methods in the website to allow user to input their own query for processing and to then view the results</i></li><li>• <i>Implementing Bayesian Networks using Python Pomegranate</i></li></ul>	

## Research Assistant

2019

- *Correcting both conceptual assignments for AUB's graduate course Discrete Models for Differential Equations (Math 350)*
- *Correcting both MatLab assignments for AUB's graduate course Discrete Models for Differential Equations (Math 350)*
- *Correcting both conceptual assignments for AUB's Numerical Linear Algebra course (Math 281)*
- *Correcting both MatLab assignments for AUB's Numerical Linear Algebra course (Math 281)*
- *Providing specific office hours designated for students attending the courses I am involved in*

## Graduate Assistant

2018-2019

- *Correcting assignments for AUB's Functional Analysis and Partial Differential Equations course (Math 309)*
- *Correcting both MatLab exams for AUB's Numerical Computing course (Math 251)*
- *Giving the weekly recitation for AUB's Numerical Computing course (Math 251)*
- *Correcting both conceptual assignments for AUB's Numerical Linear Algebra course (Math 281)*
- *Correcting both MatLab assignments for AUB's Numerical Linear Algebra course (Math 281)*
- *Holding a weekly hour and a half session attending to student questions from all Math and Math-related courses at AUB within the "Math Clinic" initiative of AUB*
- *Providing specific office hours designated for students attending the courses I am involved in*
- *Proctored the exams of the courses I am involved in as well as proctoring many exams within the math department at AUB as well*

## PROJECTS

---

### Masters Thesis

2018-2020

- *Working on a cross-document analysis method for diagnosis extraction from electronic medical records*
- *Exploring distributional similarity measurements and Bayesian graph models to extend manually annotated data to non-annotated data*
- *Using accuracy metrics such as precision and recall to evaluate the automatic annotations my experiments are generating*
- *Enriching annotations by leveraging diagnosis graphs taken from medical textbooks*

- *Using distance metric DISCO which calculates semantic similarities between words and phrases based on the statistical analysis of large text collections to draw relations and conclusions about relevant data in medical text*
- *Building an interactive website to display my work and to allow users to get personalized results and update my results*

### Optimization Project

2019

- *I solved a company portfolio convex optimization problem for a sample investment portfolios consist of allocated amounts of money made available in some properties owned by a company.*
- *The portfolio problem objective is to maximize the expected gain at the end of all time intervals considered, while taking into consideration a number of given limitations known as constraints.*

### Optimizing Quranic Search Using SAMA and Buckwater

2018

- *An exogenesis of the Holy Quran was used to optimize any word-based Quranic search.*
- *The tokens found in the Quran and the exogenesis were stemmed using SAMA and the resulting unique instances were stored with the locations of their occurrences in the Quran and exogenesis with the respective verse and chapter number.*
- *Synonym/Antonym/Closeness relationships were determined between the words occurring in the Quran and exogenesis as well as the ontology tree to further improve result search.*

### Distributional Natural Language Processing

2018

*Implemented a distributional similarity algorithm using JAVA and R to identify key words in a medical record. Key words given by practicing physicians were used for training and testing accuracy of data.*

### CUDA HMM implementation

2018

*Used C to replicate a Hidden Markov Model in CUDA, a parallel programming environment, and applied the research data I co-own to it.*

### Clustering to Estimate

2017

*Used a statistical clustering method to estimate the number of trees on the campus of the American University of Beirut and the expected circumference.*

## SKILLS

Languages	English (mother tongue) Arabic (highly fluent)
Software	JAVA, PYTHON, HTML, JAVASCRIPT, R, ANGULAR (BASICS), GITHUB, MATLAB (BASICS), C (BASICS), C++ (BASICS)

## **REFERENCES**

---

**Professor Fadi Zaraket**

Profession: AUB CCE Professor

Relation: RA advisor

Contact:

Email: fz11@aub.edu.lb

Number: (+961) 01350000 ext 3484

**Professor Nabil Nassif**

Profession: AUB Mathematics Professor

Relation: RA and GA advisor

Contact:

Email: nn12@aub.edu.lb

Number: (+961) 01350000 ext 4227