

Time Based Non-Negative Matrix Factorization (NMF) for Emergent Topics on Social Media

Project Description: Text stream analysis becomes an important approach for discovering new topics. Automatic extraction of these new topics from huge amount of documents will help identifying events such as crisis taking place in real-time. Therefore, we will benefit from the non-negative matrix factorization (NMF) technique in discovering new topics for sudden crises.

Duties/Activities: The Intern will work on an existing baseline code for NMF and get supervision from mentors in tweaking the code to match the designated problem. Then, the intern will test the developed code on real Twitter data sets.

Required Skills: Understand the basic machine learning algorithms. Have a good programming language skills such as Python, R, MATLAB, C++, or Java.

Preferred Intern Academic Level: Bachelor degree.

Learning Opportunities: : Learn how to do a research and working with a group of scientists. Also, understand how to formulate a physical problem into mathematical equations and programming language. In addition, Learn how to use and develop non-negative matrix factorization (NMF) algorithm on real data set.

Expected Team Size: 3

Mentors

Name: Mohamed Elshrif

email: melshrif@hbku.edu.qa

Name: Muhammad Imran

email: mimran@hbku.edu.qa