ERCIM "ALAIN BENSOUSSAN"
FELLOWSHIP PROGRAMME

Scientific Report

First name / Family name	Rasul Enayatifar
Nationality	Iranian
Name of the *Host Organisation*	NTNU
First Name / family name of the *Scientific Coordinator*	Professor Pauline Catriona Haddow
Period of the fellowship	01/02/2020 to 31/08/2020

I – SCIENTIFIC ACTIVITY DURING YOUR FELLOWSHIP

My contributions for a 7-month contract with FP Research Training Programme include three main activities as follows:

1. Applying a grant
   Collaborating with Prof. Pauline Haddow and Dr. Kazi Shah Nawaz Ripon (Østfold University, Oslo) to apply for a grant related to my ERCIM project. I was in charge of leading one of the work packages. The project aims at developping bio-inspired techniques to solve drone vehicle routing problem for health care usages.

2. Attending post-graduate student meeting
   My supervisor (Prof. Haddow) had weekly meetings with her post-graduate students for the sake of brainstorming and discussing about their projects. I attended all the face-to-face and on-line meetings (after lock-down) since the beginning of my arrival to Norway.

3. Working on Atriusm project
   I had a good collaboration with Prof. Pualine Haddow and Prof. Pinar Øzturk (NTNU) to work on Atriusm project which defined by Prof. Øzturk. This project aimed to take advantage of bio-inspired techniques abilities to be applied on Atriusm concept.

II – PUBLICATION(S) DURING YOUR FELLOWSHIP

I have completed following paper during my ERCIM fellowship. It is not submitted yet.

Ensemble of Distance and Similarity Measures for Noise Detection using Learning Automata Particle Swarm Optimization

Hossein Nematzadeh [1], Rasul Enayatifar[2,*]

*** Corresponding author: rasul.enayatifar@ntnu.no**

Abstract- Noisy data can adversely affect the results of any data classification analysis through creating unaccurate prediction. In this paper, a noise detection method for two-labeled medical datasets is proposed which incorporates three stages. First, the dimension of the given dataset is reduced so that half of the trivial features are discarded using Learning Automata Particle Swarm Optimization (LAPSO) algorithm. LA tunes inertia weights and causes better convergence. Second, Euclidean Distance, Cross Mathematical Expectation, and Correlation Coefficient have been applied as measures to distinctively calculate the distance and similarity regarding both samples within a class label and samples of different class labels. Majority voting function has also been applied over three measures to construct the noise free datasets by either removing or relabeling distinguished noisy samples based on θ (θ specifies removing or relabeling ratio). Third, the final noise free data set is determined through a wrapper-based evaluation using Naïve Bayes, Support Vector Machines, and Decision Tree. The proposed method is validated through three standard and four high dimensional benchmark medical datasets. In addition, trajectory, stability, and effectiveness rate of the embedded LAPSO are evaluated and noise injection analysis was also investigated with θ^*. Extensive experiments show that the proposed noise detection method improves the prediction of classifiers in presence of borderline samples.

Keywords: Noise detection, learning automata, particle swarm optimization (PSO), Similarity measure

III – ATTENDED SEMINARS, WORKHOPS, CONFERENCES

None

IV – RESEARCH EXCHANGE PROGRAMME (REP)

None

I