

IAF SYMPOSIUM ON INTEGRATED APPLICATIONS (B5)  
Integrated Applications End-to-End Solutions (2)

Author: Mr. Saeed Al Mansoori

Mohammed Bin Rashid Space Centre (MBRSC), United Arab Emirates, saeed.almansoori@mbrsc.ae

SPATIO-TEMPORAL STATISTICAL ANALYSIS AND DEEP LEARNING TECHNIQUES FOR  
TRAFFIC ACCIDENTS PREDICTION

**Abstract**

Trac accidents impose significant problems in our daily life due to the huge social, environmental, and economic expenses associated with them. The rapid development in data science, geographic data collection, and processing methods encourage researchers to evaluate, delineate trac accident hotspots, and to effectively predict and estimate trac accidents. Moreover, the pattern and association of hot and cold spots distribution can be examined to identify features with the most spatio-temporal effect on trac accidents. In this study, Kaggle trac accidents dataset that covers United Kingdom for the time period between 2012-2014, with more than four hundred thousand event associated with twelve features, will be investigated. These features consist of location, time, month, day of week, weather conditions, light conditions, road surface, road type, and area (urban or rural). The methodology consists of three main techniques. First, Moran's I method of spatial autocorrelation, and Getis-Ord  $G_i^*$  statistics will be used to examine and relate trac accidents dataset in terms of spatial and temporal features. The spatial statistics and cluster investigation will result in accident hot and cold spots. Feature weights will be assigned according to statistically significant distribution of accident features along with Z-score values. Second, weighted features will be used as inputs for a Deep FFNN (DFFNN). For this purpose, sigmoid is used as an activation function. Additionally, the same inputs will also be tested using SVM in order to compare the results at a later stage. Finally, the performance of the proposed DFFNN will be evaluated based on its accuracy, misclassification rate, precision, prevalence, histogram of errors, and confusion matrix.