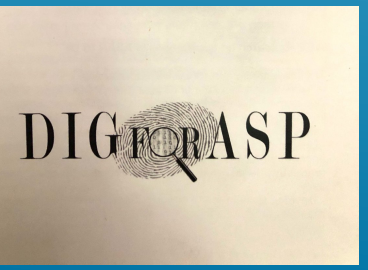
NEURAL NETWORKS FOR SAN FRANCISCO CRIME DATA PREDICTION

Lorenzo De Lauretis and Stefania Costantini

Digital Forensics Evidence Analysis via Intelligent Systems and Practices



UNIVERSITÀ
DEGLI STUDI
DE L'AQUILA

MACHINE LEARNING

- Machine learning is an application of Artificial Intelligence, that provides systems with the ability to automatically learn and improve from experience without being explicitly programmed
- During the past few years, machine learning became ever more popular
- It can be used to simplify everyday life and is applicable to a lot of scenarios

OUR WORK

- In our work, we have examined, through machine learning, a dataset concerning San Francisco crime data, from January 2003 to March 2015
- The dataset is represented using a CSV file

THE DATASET

- The dataset have the following structure:
- **DATES**
- CATEGORY
- DESCRIPT
- DAYOFWEEK
- PDISTRICT
- RESOLUTION
- ADDRESS
- **X**
- **Y**

SOFTWARE

- The software we used to create the neural net and elaborate the data was RapidMiner
- The neural net provides us with useful statistical data about the crimes committed in the districts of San Francisco

- Applying machine learning techniques via our neural net, we will be able to make classification on the crime data
- We used Deep Learning Model
- In particular, we can discover whenever there is an increase/ decrease of a particular kind of crime in certain areas

MACHINE LEARNING RUNNING

THE PREDICTION

- Using RapidMiner, we can predict the major crime that will happen in a particular day of the week of the selected date
- Having those data as inputs:
  - PDistrict
  - Dates
  - DayOfWeek
- We will have a prediction of the crime that potentially will be committed in that day in that district

- We choose as Date from our simulation the wednesday 28/10/2020
- The Police District we choose is the Northern one

- From the simulation, in Northern district, on wednsdays 28/10/2020, there will be an increase of «Grand theft of locked auto»
- The Northern district can assign more policemen to this specific task

EXAMPLE

OBJECTIVES

- Our work can help the police districts for the assignment task of policemen, cars and other resources, increasing the attention on the districts with an increasing crime type
- If, according to our neural network, in a particular day of the year, a particular kind of crime is going to increase in a particular district, more specialized policemen should be assigned

DISCUSSION

- By using a neural network, we are able to make previsions on the crime rate
- We are not able to know the reasons underlying the increase or decrease of crime rate and the real happenings of crime events
- In a previous work we analyzed London crime data of the same period. That dataset lacked more specific information such as crime geolocalization and Date. With this specific dataset, we have more information that allows us to do more precise previsions
- The dataset that we have analyzed unfortunately lacks more specific information, such as **the profiles of perpetrators and victims**, features that would have made our results more interesting

FUTURE WORKS

- To tackle the problems we saw in “Discussion”, we propose for future work the adoption of Computational Logic
- In particular, we will experiment the adoption of Inductive Logic Programming which is a form of Machine Learning that however learns rules, that should represent causal connections extracted from data to understand “why” they increase/decrease in each specific area, assuming that provided data are richer than those we have examined
- In complement, forms of reasoning such as Answer Set Programming might elicit future plausible scenarios of crime distribution.

THANKS FOR YOUR ATTENTION ☺