

A FOG COMPUTING PROTOTYPE

Course Project for Big Data Analytics — Winter 2019

Ali Alizadeh Mansouri Marco Sassano

April 15, 2019

Concordia University

TABLE OF CONTENTS

Introduction

Problem

Methodology

Edge Computing

Cloud Computing

Conclusion

INTRODUCTION

WHAT IS FOG COMPUTING?

Internet of Things (IoT)

- 50 billion devices (sensors, smart phones, smart cities, healthcare, smart vehicles, ...)
- 100s of terabytes towards petabytes per day
- limited computing resources

WHAT IS FOG COMPUTING?

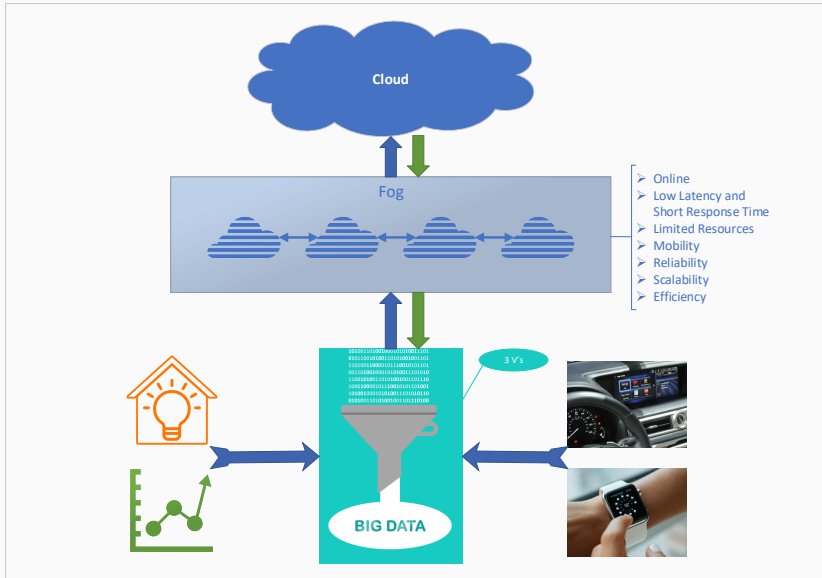
Internet of Things (IoT)

- 50 billion devices (sensors, smart phones, smart cities, healthcare, smart vehicles, ...)
- 100s of terabytes towards petabytes per day
- limited computing resources

Cloud computing

- Flexible economy
- Scalability
- Adaptation to varieties of computational demands
- Suffers from high latency

A FOG COMPUTING ENVIRONMENT



PROBLEM

Early detection of epilepsy seizures using EEG timeseries data

- Stream processing on the edge
- Clustering in the cloud

DATASET

- 500 individuals, each with 4097 data points for 23.5s
- $23 \times 500 = 11500$ records, each record contains 178 data points (columns) for 1 second
- Restructured to (time, value) tuples to be used as a stream
 $\Rightarrow 11500 \times 178 = 2\,047\,000$ tuples (windows of 178)
- 5 classes \rightarrow 2 classes (binary)

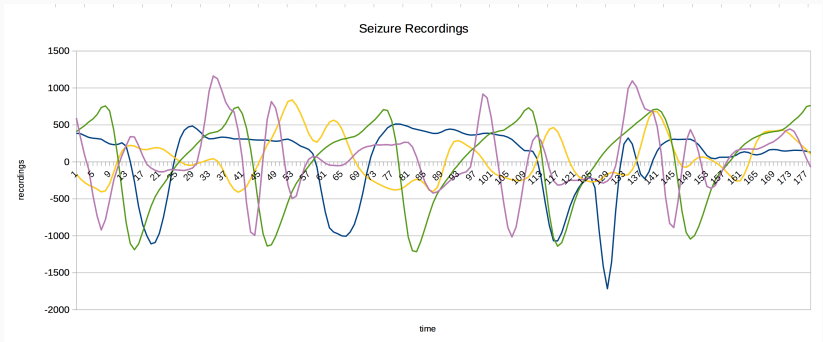
	Total	Positive	Negative
Training	7 000	1 392	5 608
Test	4 500	908	3 592
	11 500	2 300	9 200

METHODOLOGY

STREAM FILTERING

Goal Filter the stream for out-of-range tuples/anomalies

- light-weight on resources
- fast



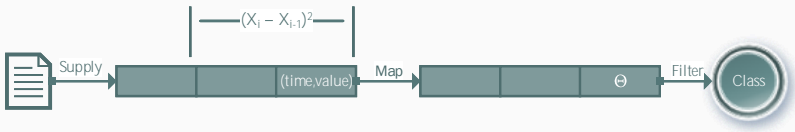
Different statistical measures: variance, skewness, kurtosis

Selected measure

$$\Theta = \sqrt{\sum_{i=2}^n (X_i - X_{i-1})^2}$$

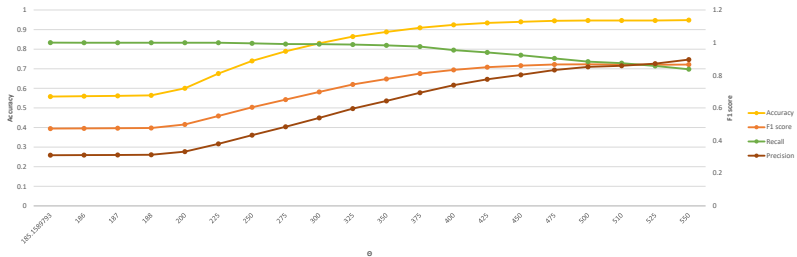
Apache Edgent: Programming model and runtime

- **Supply** reads the dataset and supplies it as a stream
- **Map** maps the tuples to the selected measure
- **Analysis** classifies the out-of-range Θ values as seizures

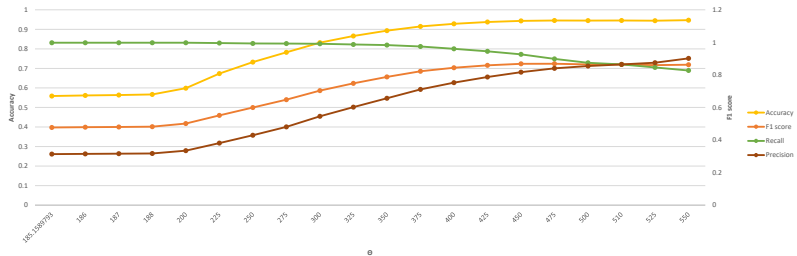


RESULTS

TRAINING DATA



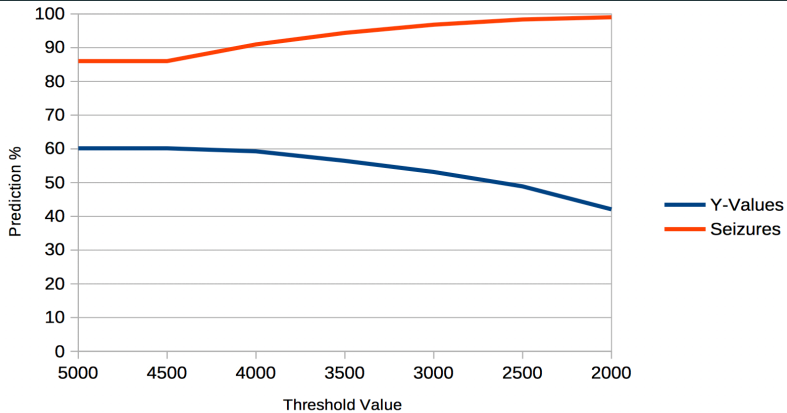
TEST DATA



Selected measure

$$s = \sum_{i=2}^n |X_i - X_{i-1}|$$

RESULTS



CONCLUSION

- Tests with more than one sensor
- Introducing a publish-subscribe platform such as Apache Kafka, IBM Watson IoT Hub, etc. (e.g. using MQTT as the protocol)

Thank you!