Towards an On-board Personal Data Mining Framework For P4 Medicine

Dr. Mohamed Boukhebouze
Deputy Department Manager, CETIC

European Data Forum 2015, November 16-17
Luxembourg
Agenda

- Overview of CETIC

- On-board personal data mining framework
  - Introduction
  - Objective & Challenges
  - Architecture Overview

- Use case: Epilepsy seizure

- Discussion
Agenda

• Overview of CETIC

• On-board personal data mining framework
  • Introduction
  • Objective & Challenges
  • Architecture Overview

• Use case: Epilepsy seizure

• Discussion
CETIC - Overview

Created in 2001
- Accredited Research Centre
- ~40 researchers in 3 departments

Mission
- Applied Research at EU and Regional level
- Technology Transfert Agent to (inter-)Regional Industry

Serving Industry
- SotA/Techno Evaluation and Coaching (HW/SW)
- Trusted Third Party
- Connect Industry to latest research results

International Involvement
- EU: FP7 - H2020 – Coordinator and Participants
- EraNets and Interreg
- Regional Research project (Plan Marshall 2.vert)

Regional Funding for Contract Research
- « Technological Checks »
- Feasibility Studies
- R&D 1-1 Projects
CETIC expertises

SOFTWARE & SYSTEM ENGINEERING

PROCESS & PRODUCT QUALITY
OPTIMISATION
MODELLING
SECURITY

SOFTWARE & SERVICES TECHNOLOGIES

FUTURE INTERNET
CLOUD COMPUTING
SEMANTIC WEB
BIG DATA
OPEN SOURCE & OPEN DATA

EMBEDDED & COMMUNICATING SYSTEMS

INTERNET OF THINGS
PROGRAMMABLE SYSTEMS
SMART OBJECTS
WIRELESS SENSOR NETWORK
CETIC Experience in Big Data

**SCALABILITY**
- Volume
- Variety
- Velocity

**Data Acquisition & Integration**
- APIs, Web crawling
- ETL, Semantic analyzers,

**Storage**
- Benchmarking & taxonomy of storage solutions
- Data infrastructure implementation

**Pre-processing & Querying**
- Linked Data (RDF)
- Ontologies conception
- Data quality

**Real Time Analysis**
- IoT On-Board Predictive, Personalised and Preventive Analysis

**Data Analysis**
- Process Mining
- Text Analysis
- Data Mining

**EMBEDDED & COMMUNICATING SYSTEMS**
Agenda

• Overview of CETIC

• On-board personal data mining framework
  • Introduction
  • Objective & Challenges
  • Architecture Overview

• Use case: Epilepsy seizure

• Discussion
On-board data mining framework: Introduction

• Wearable devices can play a major role in the ubiquitous collection of a large subset of personal data
  • Physiological sensors (e.g. EEG, ECG),
  • Environmental sensors (e.g. temperature monitor),
  • Location sensors (e.g. GPS)

• Wearable devices are used as interface of data mining algorithms
  • Execution on a high performance computational facility (remote processing)
    • Like: a cloud-based infrastructure

• Issue
  • In several scenarios, ubiquitous and offline data stream mining is required
    • So that, a continuous monitoring and real-time processing can be done

Use wearable devices not only to collect personal data, but also to on-board execute the personal data mining algorithms (local processing)
On-board data mining framework: Objective

- Development of an on-board personal data-mining framework for P4 medicine:
  - **Prediction**
    - Predicting the risk of developing diseases or faintness
  - **Prevention (Prescription)**
    - Providing recommendations that help to avoid or reduce the risk of developing certain diseases or faintness
  - **Personalisation**
    - Personalising the prediction and prevention recommendation by taking into account the patient’s context
  - **Participation**
    - Participation of the patients in the learning and tuning of the prediction and prevention models by providing feedback about issued notifications and recommendations.
On-board data mining framework: Challenges

- **Resources limitation**
  - Wearable devices are limited in terms of computational power, energy consumption, storage / memory capacities and bandwidth

- **Context and resources changes**
  - The ubiquitous data stream mining is faced with the problem of patient situation changes and devices disconnection

- **Data quality**
  - Despite the fact that current sensors are becoming more sophisticated, accuracy of gathered data is not always ensured
On-board data mining framework: Solutions

- The framework deals with on-board processing challenges by combining three approaches:
  - **A distributed data mining approach**
    - Execute data mining algorithms concurrently over the devices network
  - **A context-aware and resource-aware adaptation approach**
    - Detect changes of user situation or devices and adjust automatically parameters of data mining algorithms
    - Adjust dynamically load distribution of the data mining algorithms based on resources availability
  - **A probabilistic data mining approach**
    - Compute the data variation and uncertainty based on a data quality model
    - Define a mechanism that allows marking the data as “uncertain” to prevent its further analytical usage
Architecture Overview

Web services

User profile
Device profile
Environment
Context Model

Data streaming engine

Preventive Analysis
Personalisation
Predictive Analysis
Data Analysis

Buffer

Risk disease Alarm
Recommendation avoid/reduce disease risks

Cloud based data processing & Storage

Node (sensor)

Buffer
Data Analysis

Data
Command

Node (sensor)

Buffer
Data Analysis

Data
Command

Node (sensor)

Buffer
Data Analysis

Data
Command

Master Node

Risk disease Patterns
Data Quality Model
Agenda

• Overview of CETIC

• On-board personal data mining framework
  • Introduction
  • Objective & Challenges
  • Architecture Overview

• Use case: Epilepsy seizure

• Discussion
Use Cases: Refractory Epilepsy (Introduction)

- The refractory epilepsy patients do not respond to conventional treatments (30% of all patients)
  - It has a strong impact on quality of life since the seizures are random and not controlled with medications

- The proposed framework can be used to continuously monitor refractory epilepsy patients:
  - Early predicting seizures based on the patterns that are developed by our partner ULB, Belgium
  - Refining the prediction pattern based on the context of the patient
  - Providing context-aware based recommendations to avoid the seizures risk by analysing the factors that favour their appearance (e.g. stress level)
Architecture: Epilepsy Seizure Use Case

Web services

- User profile
- Device profile
- Environment
- Context Model

Data streaming engine

- Preventive Analysis
- Personalisation
- Predictive Analysis
- Data streaming engine

Master Node

- Epilepsy Seizure Patterns'

Feedback

Risk seizure Alarm

Recommendation avoid/reduce seizure risks

Cloud based data processing & Storage

Command

Data

EEG

Stress sensor

Activity tracker

www.cetic.be
Use Cases: Refractory Epilepsy  (Related Work)

- Embrace Watch (from Empatica, Italy)
  - Seizure monitoring (detection and notification)
  - Activity Tracking
  - Stress Management
  - Sleep Monitoring

Weakness

- No seizure prediction
- No seizure prevention
Use Cases: Refractory Epilepsy (Related Work)

- BioSerenity

Weakness

- No seizure prediction
- No seizure prevention
Use Cases: Refractory Epilepsy (Related Work)

- NeuroPro (Swiss-based medical technology company)
  - Development of user-friendly and wireless EEG headset
  - The EEG data is then processed in real-time in a server
  - A mobile application can report and alert the user with the results in real-time

Weakness

- No seizure prevention
Agenda

- Overview of CETIC

- On-board personal data mining framework
  - Introduction
  - Objective & Challenges
  - Architecture Overview

- Use case: Epilepsy seizure

- Discussion
Discussion

• The proposed framework can be applied in many others uses cases
  • e.g. Heart attack, Human Falls, ...

• The collected data can be used to improve the disease patterns as well as the healing treatment

• Open issues
  • What about acceptance of the wearable devices by the patients?
  • What about precision of the disease patterns?
  • What about data privacy?
Thank you for your attention

Dr. Mohamed Boukhebouze
Mohamed.boukhebouze@ctic.be