[El-Sappagh et al 2018] Shaker El-Sappagh, Jose M. Alonso, Farman Ali, Amjad Ali, Jun-Hyeog Jang and Kyung-Sup Kwak (2018). **An ontology-based interpretable fuzzy decision support system for diabetes diagnosis.** *IEEE Access, 6:37371-37394.*

An ontology-based interpretable fuzzy decision support system for diabetes diagnosis

Alejandro Ariza Casabona

14/10/2019

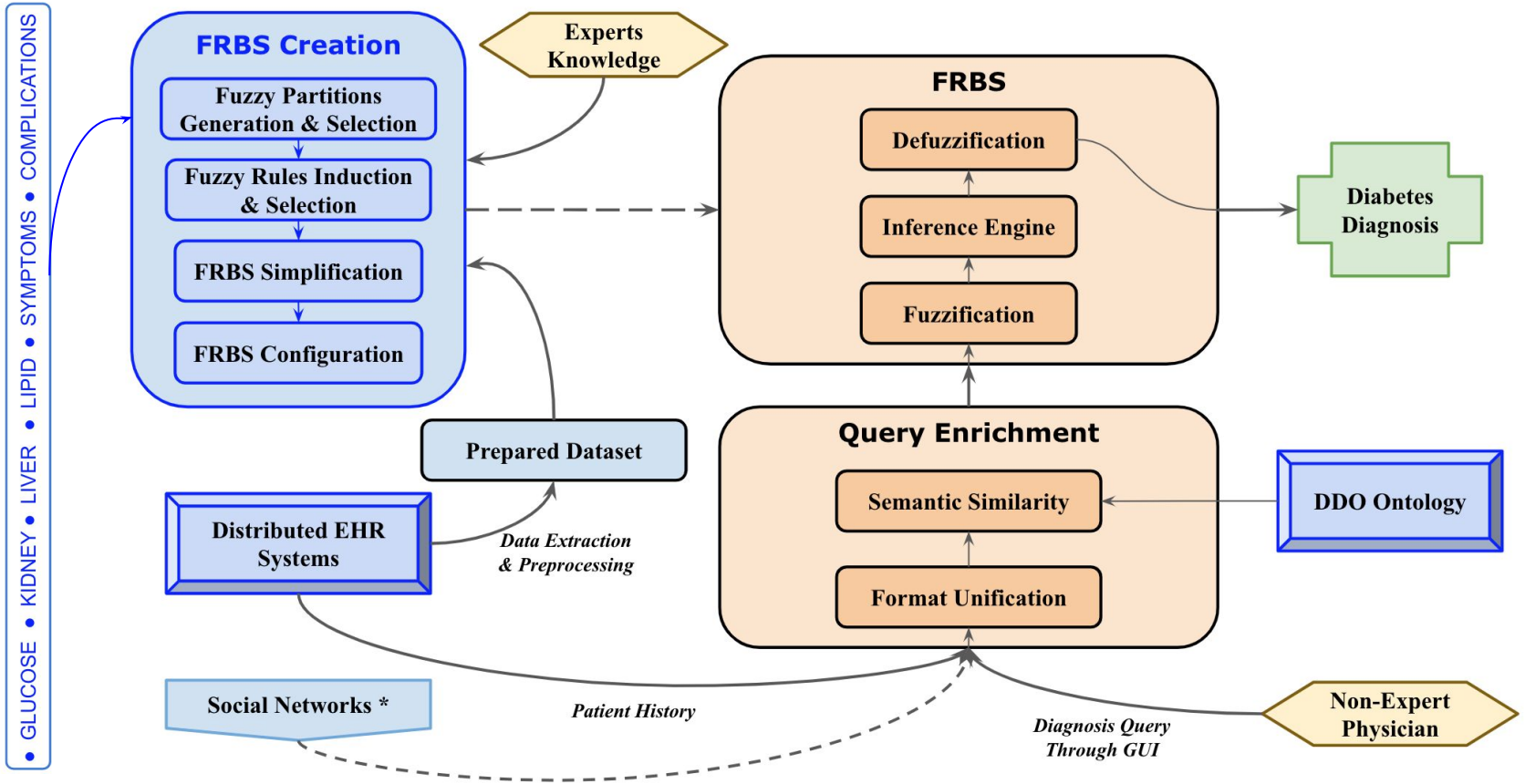
UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

OVERVIEW

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

- ***Problem Domain***: Clinical Decision Support System (CDSS) for **Diabetes** Mellitus (DM) **Diagnosis**
- ***Goals***:
  - Great **accuracy** in the DM Diagnosis.
  - Highly **interpretable and dynamic** Fuzzy Rules and Fuzzy Partitions (medically intuitive).
  - **Semantic Reasoning** to enable interoperability with Electronic Health Records (EHR) systems.
- ***Data Sources***:
  - Evaluation Dataset obtained from the **Mansoura University Hospitals (Egypt)** from January 2010 to August 2013.
  - **Distributed EHR Systems** (Patient History) and Experts Knowledge (Interpretability).
  - Connection to **Social Networks** (Future Work).
- ***IDSS Scope***:
  - Build an accurate and semantically interpretable hierarchical FRBS that supports the DM diagnosis of non-expert physician's (or patient's) queries (including the creation of a GUI interface).
  - Address the possible isolation of subsystems in the decision making process.
  - Add **ontology reasoning** to improve the semantic expressiveness and interoperability of the CDSS

ARCHITECTURE OF THE IDSS



IDSS METHODOLOGY

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

- ***Preprocessing Techniques***:
  - Outliers detection, missing data handling, data transformation (categorical features encoding and semantic encoding of medical concepts according to SCT) and feature selection.
- ***Fuzzy Rule Base Systems (FRBS) Creation***:
  - Fuzzy partitioning: K-Means Clustering or Hierarchical Fuzzy Partitioning (HFP) based on the Partition Coefficient (PC), the Partition Entropy (PE) and the Chen Index (CI).
  - Fuzzy Rules Induction: Fuzzy Decision Trees (FDT) or Wang and Mendel (WM) methods.
  - Rules Weighting: Fuzzy Analytical Hierarchy Process (FAHP).
  - FRBS Simplification: Based on HILK++.
  - FRBS Configuration and Implementation: Mamdani Inference engine and “Mean of Max” defuzzification.
- ***FRBS Semantic Extension***:
  - Semantic Similarity based on DDO Ontology using the Pellet reasoner and the SNOMED CT Terminology.

IDSS EVALUATION

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

- ***Comparison among different versions*** of the model for each sub-FRBS:
  - {3, 5, 7} Fuzzy partitions (different clustering algorithms e.g. uniform, K-Means and HFP).
  - WM or FDT.
  - With or without FRBS Simplification.
  - Weighted Hierarchical FRBS (with FAHP).
  - Semantically intelligent WH-FRBS.
  - SWH-FRBS after isolating specific Subsystems.
- ***Comparison against other ML techniques***: naïve Bayes, support vector machine (SVM), logistic regression, k-nearest neighbor (KNN) with k = 3, decision tree based on C4.5, artificial neural network (ANN) and random forest.
- ***Evaluation Metrics (averaged over 10-fold cross-validation)***:
  - Quality metrics: Accuracy, Precision, Recall and F-Measure.
  - Interpretability Indexes: Rules and Premises.
  - Partition Quality Indexes: PC, PE and CI.

GOOD ACCURACY

MEDICALLY INTUITIVE