TOPIC: Real-World Data & Information Systems – Personalized & Precision Medicine

Acceptance Code: RWD127

PRELIMINARY RESULTS OF A MACHINE LEARNING MODEL FOR THE DETECTION OF UNDIAGNOSED CASES OF ANKYLOSING SPONDYLITIS IN ITALY

Diletta Valsecchi[1], Federica Moscatelli Spinelli[1], Enrico Quattrina[1], Valentina Perrone[2], Biagio Iacolare[2], Luca Degli Esposti[2], Aurelio Pio Nardozza[1]

[1]Novartis Farma S.p.A., Milan, Italy; [2]CliCon S.r.l. Società Benefit Health, Economics & Outcomes Research, Bologna, Italia

CliCon **CLICON S.R.L. SOCIETÀ BENEFIT,** HEALTH, ECONOMICS & OUTCOMES RESEARCH

ISPOR Europe 2024, Barcelona (Spain), 17-20 November 2024

BACKGROUND AND OBJECTIVES

- Rapidly evolving and disabling diseases might significantly benefit from early diagnosis. Novel technological tools have great potential applications to decrease the number of undiagnosed cases, thus allowing prompt therapy interventions [1].
- The advent of Artificial Intelligence (AI) is paving the way for new medical avenues, diagnostic options and therapeutic practices [2,3].
- AI technologies such as machine learning (ML) and deep learning (DL) might represent a powerful tool for healthcare providers to overcome the major challenges associated with timely disease detection, management, accessibility, and treatment optimization.
- For underdiagnosed conditions, important advancements might be implemented by using historical clinical data to design an algorithm for the early identification of subjects potentially affected, but not yet diagnosed, and ultimately to define a panel of disease predictors.

OBJECTIVES. This is a preliminary model in the setting of Ankylosing Spondylitis (AS) based on a synergistic utilization of ML and patient healthcare data available in administrative databases. The model was conceived to provide the clinicians with a forerunner template applicable to other diseases, where timely detection of undiagnosed cases is critical in medical decision-making.

PATIENTS AND METHODS

DATA SOURCE

A dedicated algorithm was developed using data from administrative databases of Italian healthcare entities covering about 12 million health-assisted subjects.

STUDY POPULATION: Inclusion criteria

From January 2009 to September 2023, patients with a diagnosis of AS were identified by the presence of:

- at least one hospitalization for AS (ICD-9-CM: 720); **OR**
- at least exemption for AS (codes: 054)

ML MODELS

Drug prescriptions, hospitalization discharge diagnoses and exemption codes extracted from Jan-2009 to Sep-2023 were the input variables to deploy two ML models, based on a **neural network** approach, that is a system to mimic the human brain (**Fig. 1**) and a **random forest** algorithm to provide a prediction model as output (**Fig. 2**).

NEURAL NETWORK MODEL

Neural networks are computing systems composed of layers of interconnected units. Data travel from the first layer (the **input layer**), passing through multiple intermediate layers (**hidden layers**), to the last layer (the **output layer**). The output layer produces the final result of the classification/regression task.

The neural network was built as follows:

- **Input layer**: 1181 nodes (one for each patient feature → 4th level ATC code, ICD9 code)
- **2 hidden layers**: 1500 nodes
- **Output layer**: 2 nodes (one for each outcome → with AS/without AS)

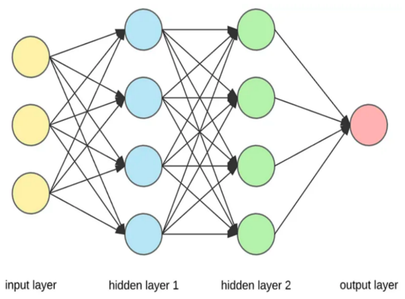


Figure 1. Neural network architecture

RANDOM FOREST MODEL

Ensemble learning algorithm for both classification and regression, that is used to combine the outputs of multiple decision trees to provide the final result. For classification task, the final outcome is the one returned by most decision trees, while for regression tasks the result is usually computed as the average between values returned by all decision trees.

Each decision tree is an individual predictive model, that consist of multiple decision nodes. Each node split data based on values of features until reaching the leaf node. A maximum depth parameter can be selected to limit the overfitting of the model and consider just the most relevant features.

The random forest was built as follows:

- **Number of decision trees**: 150
- **Maximum depth:** 50

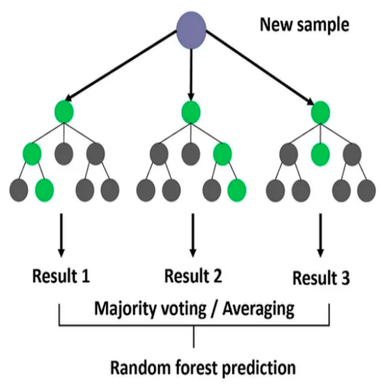


Figure 2. Random forest architecture

FUNDING: *The report used for this poster was financed by Novartis Farma S.p.A., Italy*

RESULTS

IDENTIFICATION OF THE STUDY POPULATION

From around 12 million citizens, 12,294 patients with AS were identified in the database during the inclusion period. Only adult patients with at least 12 months of data availability before inclusion (date of the first record of an AS-related exemption code or hospitalization) were selected (n=7,934): mean age was 57 years, and 46% males.

A matched control population was also selected to build the models.

CONFUSION MATRICES

Confusion matrices are frequently used in ML to represent the number of true positive, true negative, false positive and false negative predictions. Each column represents the predicted class, while each row represents the observed class:

- top left -> true negatives
- top right -> false positives
- bottom left -> false negatives
- bottom right -> true positives

Both models showed a **good ability to distinguish between AS and non-AS patients** (**Fig. 3**).

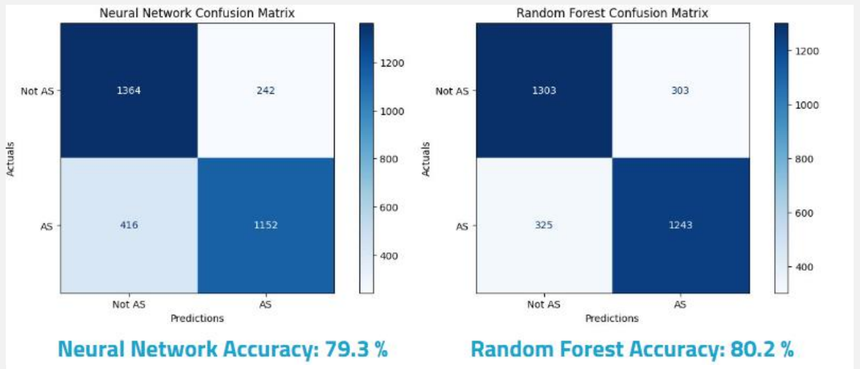
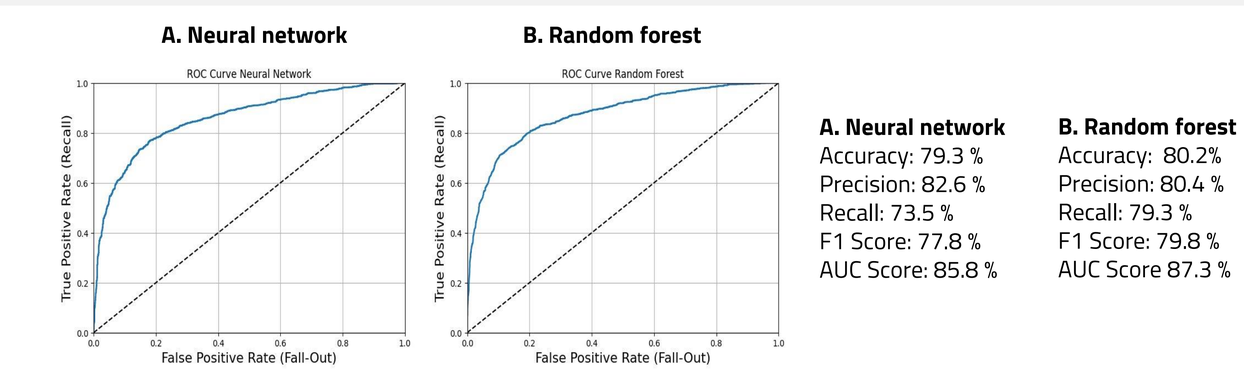


Figure 3. Accuracy score (percentage of correct predictions over all perfomed predictions) of the two models

MODELS PERFORMANCE

The following performance metrics were computed to evaluate model performance:

- **Accuracy**: percentage of correct predictions over all performed predictions.
- **Precision**: percentage of true positive predictions over all positive predictions.
- **Recall**: percentage of true positive predictions over all positive samples.
- **F1 score**: harmonic mean of precision and recall.
- **AUC score**: evaluates the area under the Receiver Operating Characteristics (ROC) curve
- **ROC curve**: plot of the True Positive Rate (TPR) against the False Positive Rate (FPR) as the classification threshold is varied. It displays the trade-off between sensitivity (how well the classifier detects true positives) and specificity (how well it avoids false positives).



A. Neural network
Accuracy: 79.3 %
Precision: 82.6 %
Recall: 73.5 %
F1 Score: 77.8 %
AUC Score: 85.8 %

B. Random forest
Accuracy: 80.2%
Precision: 80.4 %
Recall: 79.3 %
F1 Score: 79.8 %
AUC Score 87.3 %

Figure 4. Performance metrics computed for both **neural network** and **random forest** models

MAIN PROGNOSTIC FACTORS

Fig. 5 describes the **20 most important features to detect AS diagnosis**: the strongest predictors of AS were prescriptions for autoimmune diseases, hospitalization code ICD-9-CM 724 (other and unspecified back disorders), and exemption code 009 (specific for inflammatory bowel diseases - IBD).

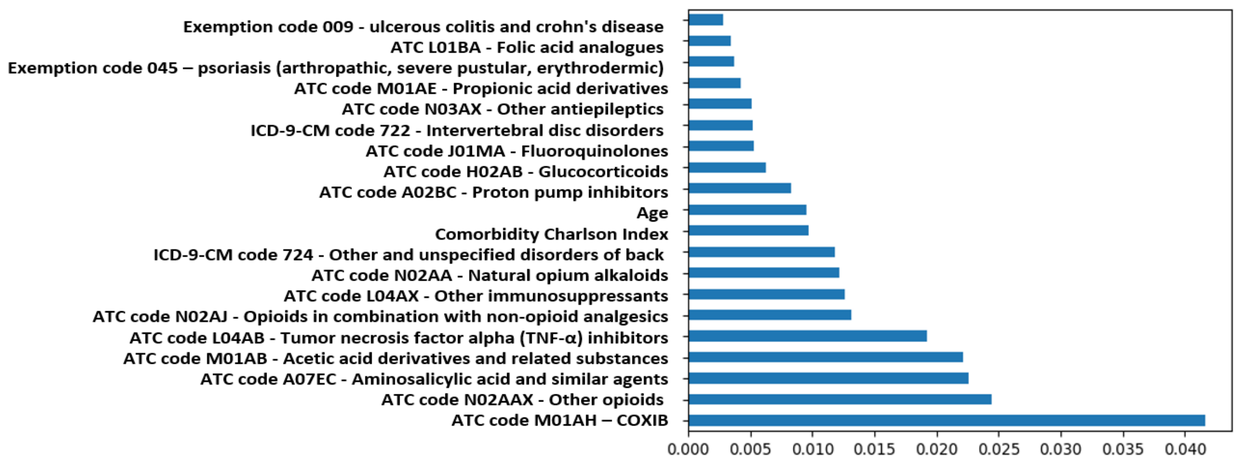


Figure 5. Top-20 most important features to detect AS diagnosis

REFERENCES

1. Iqbal J, et al. Reimagining Healthcare: Unleashing the Power of Artificial Intelligence in Medicine. Cureus. 2023 Sep 4;15(9):e44658.
2. Ravindranath R, et al. The Impact of Race, Ethnicity, and Sex on Fairness in Artificial Intelligence for Glaucoma Prediction Models. Ophthalmol Sci. 2024;5(1):100596.
3. Lee SH, et al. Development of Clinically Validated Artificial Intelligence Model for Detecting ST-segment Elevation Myocardial Infarction. Ann Emerg Med. 2024;84(5):540-548.

CONCLUSIONS

- The present model describes a successful joint utilization of ML and real-world data applied to the setting of AS on a sample of Italian health-assisted citizens corresponding to nearly 20% of the country population.
- This pivotal design might pave the way for further clinical applications, providing a valuable tool to facilitate the early detection of undiagnosed cases of special importance for rapidly evolving and disabling diseases.