University Grenoble Alpes
L3/EAAD
DALLAORANE WILFRED

Academic Year: 2023-2024
Mathematics 6 Individual Report

Identification of Age-Related Conditions

Detecting Health Issues Using Anonymized Data and Artificial Intelligence

1 Context

Aging can lead to multiple health issues such as heart disease, dementia, hearing loss, and arthritis. Therefore, developing effective methods to detect these conditions is essential. Before artificial intelligence (AI) became widespread, predicting age-related conditions relied primarily on methods involving data collection from patients. These methods had limited performance because they were often time-consuming, intrusive, and could undesirably disrupt patient privacy, requiring a significant amount of data.

According to the Larousse dictionary, artificial intelligence (AI) is perceived as the ability of machines to learn and act like humans by using instructions and data. It helps solve problems and make decisions. Yann LeCun defines AI simply as: *"enabling machines to perform tasks that are usually assigned to animals and humans"*. In short, AI is a vast field that touches upon computing, mathematics, neuroscience, and even general health.

AI brings a revolutionary approach to analyzing data in-depth. Unlike traditional methods, predictive models like **XGBoost** or **random forest** are capable of learning complex patterns from key information related to conditions. They identify relationships between health features and age-related conditions by using various ways to improve their predictions, helping to predict more accurately whether an individual is affected. These models also shorten the process and preserve patient data privacy but are limited when dealing with critical issues where lives are at stake. AI works by collecting key features related to conditions and encoding them.

2 Technical Description of AI

The available data includes more than 50 anonymized health features, primarily numeric, except for one categorical variable. A neural network is a computational model inspired by the human brain. It consists of neurons that process input information and pass the output to other neurons. In this scenario, eight (08) neurons are used for each health feature. With fifty training data points, the neural network can have several hundred parameters to adjust. The output of the neural network is the probability for each class, quantifying the uncertainty of whether an

individual is affected or not. This estimates the margin of error for health characteristics to better understand its reliability.

3 Advantages and Limitations of AI

The use of AI offers significant improvements in predicting age-related conditions. It enables faster and more accurate diagnosis, which can improve treatment outcomes and reduce healthcare costs. AI aids doctors in analyzing and interpreting data related to elderly patients. In terms of efficiency, AI allows healthcare professionals to focus more on patient care by quantifying uncertainty. However, AI is not flawless. Input data may contain errors or biases, which could affect the reliability of predictions. Additionally, some conditions may require further human expertise for pre-processing or specialized analysis. For instance, if the probability of being affected is 0.49% and the probability of not being affected is 0.51%, it would be challenging to determine the patient's condition without a doctor's intervention.

While predictive models enhance the assessment of prediction reliability, they can introduce new challenges. These models require high-quality data and specialized skills for maintenance, which may entail financial costs. Moreover, data confidentiality is not guaranteed.

4 Future Outlook

The AI model proposed in this domain could transform medical practices and impact the daily lives of patients and healthcare professionals. It may raise concerns about the privacy of health data and change the dynamics of doctor-patient interactions. Yann LeCun clarifies this by stating: *"The fundamental techniques are open, meaning they are published and available. Any high school student can download the code and train a translation system. These technologies are not secret."* The generalization of AI use could pave the way for faster and more accurate diagnostic methods.

The impact on jobs may vary. Some jobs could be transformed or automated, but new specialized roles may also emerge. Ultimately, the increasing use of AI in healthcare may contribute to more personalized and efficient care but would require deep reflection on its social and ethical implications. No matter how intelligent machines become, objectivity is not their specialty, and health issues remain inherently human as they involve the communication of emotion. In the future, the goal would be to train AI on more "objective" data to make it relatively objective.

However, how do we get patients to trust the results provided by these models? Who is responsible for the use of AI? The doctor? The statistician?

This topic is inspired by the Kaggle competition [**?**].

Appendix

Function to evaluate the reliability of the prediction:

$$\log \text{loss} = \frac{-\frac{1}{N_0}\sum_{i=1}^{N_0} y_{0i} \log p_{0i} - \frac{1}{N_1}\sum_{i=1}^{N_1} y_{1i} \log p_{1i}}{2}.$$

Where:

- N_c is the number of observations in class c,
- log is the natural logarithm,
- y_{ci} is 1 if observation (i) belongs to class c, and 0 otherwise,
- p_{ci} is the predicted probability that observation i belongs to class c.

References

https://manifesto-21.com/rever-intelligence-artificielle-mieux-la-comprendre-avec-yann-lecun/
https://www.lemonde.fr/festival/article/2017/09/23/yann-lecun-de-facebook-l-intelligence-artificielle-va-sauver-des-vies-mais-il-y-a-aussi-des-dangers$_5$190311$_4$415198.*html*