Case Study

Toulouse Business School

MSc AIBA (Part Time)

July 2023 / 2024

Duration: 4h

Instructions: Any generative AI is forbidden. Use the software of your convenience.

Context

Hyperthyroidism stands as a prevalent chronic condition in Europe, impacting many individuals annually and posing a substantial economic burden on countries. Hyperthyroidism, characterized by increased thirst, frequent urination, weight loss, fatigue, presents a serious health risk as it can lead to various complications, reducing both quality of life and life expectancy. The scale of hyperthyroidism in Europe is significant and warrants attention. According to data from a European Health Agency, a considerable portion of the population is affected by hyperthyroidism, highlighting the urgent need for preventive measures and effective management strategies. Hyperthyroidism prevalence varies across demographic factors. Furthermore, hyperthyroidism imposes a substantial economic burden on the European healthcare system and society. The direct and indirect costs associated with hyperthyroidism management, including medical care, lost productivity, and disability, amount to a considerable sum annually. Addressing hyperthyroidism comprehensively requires not only individual-level interventions but also systemic changes to promote healthier lifestyles and improve access to quality healthcare services across all segments of the population.

The hospital Pitié Salpêtrière has collected data on 20000 individuals to study factors influencing hyperthyroidism disease. They want to fit a model able to predict that someone will be positive to hyperthyroidism.

Objective

The aim of this Case Study is to propose to the hospital Pitié Salpêtrière a model able to predict that someone will be positive to hyperthyroidism based on observed attributes.

Dataset

Each student will receive 01 dataset through c@mpus, containing 20000 data points stored as rows with 21 features in columns. **Data are not the same from one student to another one**.

Attribute information:

High_Blood_Pression: Have you been told that you have high blood pressure? Yes (1) or No (0)

High_Cholesterol: Have you been told that you have high Cholesterol? Yes or No

Cholesterol_Check: Did you check your Cholesterol within past five years? Yes or No

Body_Mass_Index: Calculated

Smoking: Have you smoked at least 100 cigarettes in your entire life? Yes or No

Stroke: Ever told you had a stroke ? Yes or No

Heart_Disease: Ever told you had a Heart Disease? Yes or No

Physical_Activity: Did you have physical activity or exercise during the past 30 days other than your regular job, Yes or No

Fruits: Do you consume Fruit 1 or more times per day? Yes or No

Vegetables: Do you consume vegetables 1 or more times per day? Yes or No

Alcohol: Do you have more than 14 drinks per week (for men) and more than 7 drinks per week (for women)? Yes or No

Health_care: Do you have any kind of health care coverage? Yes or No

General_Health: Rate your health from 1 to 5?

Mental_Health: How many days during the past 30 days was your mental health bad?

Physical_Health: : How many days during the past 30 days was your physical health bad?

Diff_Walk: Do you have difficulty climbing stairs? Yes or No

Sex: Sex of respondent

Age: Age category

Education: Highest grade or year of school completed

Income: Category of income

Output variable:

hyperthroidism: Diagnostic, 0 is for no hyperthyroidism and 1 is for hyperthyroidism

Evaluation criteria

Your submission report should contain clear and well explained codes and their outputs. You should present clearly and briefly your:

- Methodology
- Data description
- Modelization steps
- Model Evaluation (Performance, using error measures)
- Interpretation of results
- Recommendations / Insights

Your report will be submitted via c@mpus under the section "EXAMEN AIBA PT 2024" **before 05:30 pm.**