**INTRODUCTION MACHINE LEARNING**

**EXERCISE 6**

A red sign with white text

Description automatically generated

|  |
| --- |
| **TEACHER:** |
| Johannes Kiesel |

|  |
| --- |
| **GROUP:** |
| Group 16 |
|  |
| **SUBMITTED BY:** |
| Aaron Perez Herrera |
| Cesar Fernando Gamba Tiusaba |
| Chun Ting Lin |
| Olubunmi Emmanuel Ogunleye |

Exercise 1 : Probability Basics (1 Points)

Which of the following statements are true?

* According to the Kolmogorov axioms the statement P(A) - P(A) = 0 holds.
* A function that fulfills the Kolmogorov axioms is a probability measure.
* Two events are statistically independent P(An B) = P(A) + P(B).
* Each subset A of a sample space 12 is an event.

The true statements are.

* A function that fulfills the Kolmogorov axioms is a probability measure. (true)
* Each subset A of a sample space 1 is an event. (true)

Exercise 3 : Bayes' Rule (2+3=5 Points)

A hospital database contains diagnoses (C1 ... C5) for 8 patients along with binary observations of symptoms S1 ... S9:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Patient** | **Diagnosis** | **Symptoms** | | | | | | | | |
| **S1** | **S2** | **S3** | **S4** | **S5** | **S6** | **S7** | **S8** | **S9** |
| **1** | **C1** | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| **2** | **C2** | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 |
| **3** | **C3** | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| **4** | **C4** | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 |
| **5** | **C3** | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| **6** | **C5** | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| **7** | **C3** | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| **8** | **C2** | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |

1. Compute based on the database the prior probabilities P(Ci) for each diagnosis.

The total number of patients is 8.

* .The number of patients diagnosed with C1 is 1, so P(C1) = 1/8 = 0.125.
* The number of patients diagnosed with C2 is 2, so P(C2) = 2/8 = 0.25.
* The number of patients diagnosed with C3 is 3, so P(C3) = 3/8 = 0.375.
* The number of patients diagnosed with C4 is 1, so P(C4) = 1/8 =0.125.
* The number of patients diagnosed with C5 is 1, so P(C5) = 1/8 = 0.125.

1. Compute based on the database the posterior probabilities P(Ci | S4) for each diagnosis.

Posterior probabilities P(Ci|S4) of the diagnoses Ci given symptom S4 are calculated as follows:

The posterior probability of Ai given B is given by,

Hence, first P(B|Ai) is to be calculated.

Therefore, the posterior probability of S4 given C is calculated as:

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |

The next probabilities