**CM0669 Machine Learning and Computer Vision**

**Homework 2:** Bayesian classification.

A crime was committed in a small village. Because the criminal’s face was covered, CCTV cameras could not disclose his/her identity. Initially, investigators did not have any clues on who might have done it. However, by analysing the scene of crime, they could find a number of features which might help to efficiently narrow the search for the criminal. Indeed, the analysis reveals that the criminal was wearing white sport shoes and the shoeprints left in the scene show that the sole was highly textured. There are only ten shops selling sport shoes in the village and all of them have no more than 4 brands (Adidas, Nike, Puma, and Umbro). Investigators have used information from the shops to find the exact number of sport shoes that have been sold per brand. Furthermore, they have known the features of the sport shoes that were sold for each brand as the following tables illustrate:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Brand | Adidas | Nike | Puma | Umbro |
| Number of shoes sold. | 820 | 540 | 340 | 300 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| sole type | Adidas | Nike | Puma | Umbro |
| **Highly textured** | 205 | 200 | 95 | 120 |
| Textured | 245 | 240 | 105 | 115 |
| Non-textured | 370 | 100 | 140 | 65 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| colour | Adidas | Nike | Puma | Umbro |
| **White** | 400 | 300 | 180 | 150 |
| Black | 330 | 200 | 120 | 90 |
| brown | 90 | 40 | 40 | 60 |

Among the four brands, the investigators have decided to focus only on the customers who bought sport shoes made by Nike and ruled out the other possibilities according to the Bayesian decision rule. Apply the Naive Bayes classifier to this problem (name and calculate the appropriate probabilities to show that the suspect is very likely to have been wearing sport shoes of Nike during the crime).