

## CE 5319 Machine Learning for Civil Engineers Exercise Set 3

### Exercises

1. Finger-printing radioactive fallout.

#### Background

Consider the task of “fingerprinting” the origin of a nuclear event <sup>1</sup>. A fusion device would discharge some unreacted plutonium 238, finger-printing to either the USA, Russia, UK, France, or China. It is not believed India or North Korea have viable fusion devices. Plutonium mixed with other fallout materials is hard to detect, but even minute amounts emit alpha particles (which can be traced!).

Alpha detectors record the intensity of alpha particle strikes in counts per second. A prediction engine (regression model) is used to estimate the alpha counts (target variable) from plutonium activity (feature variable).

Consider the calibration set contained at <http://54.243.252.9/ce-5319-webroot/2-Exercises/ES-2/PLUTONIUM.DAT>. The first column is alpha strikes per second, the second is plutonium activity. These data are from four different plutonium standards containing known plutonium activity. The levels are 0.0, 5.0, 10.0, 20.0 picocuries per gram. Each standard was exposed to the detection device from 4 to 10 times and the rate of alpha strikes was observed for each replication.

#### Problem Statement

Using these data build a prediction engine that takes as input alpha strikes per second and predicts plutonium activity (an inverse regression) for a particular detector instrument.

In your model building consider:

- Do any of the data appear to be outliers?
- Likely your original prototype engine will be  $Y_\alpha \propto X_{Pu}$
- Is the variance constant at different plutonium activities?

Next consider some transformations

- First transformation  $\sqrt{Y_\alpha} \propto X_{Pu}$

---

<sup>1</sup>Something goes boom, and you have to place blame, then go to the UN security council with evidence for permission to end the source

- Second transformation  $\sqrt{Y_\alpha} \propto \sqrt{X_{Pu}}$

Using your favorite resulting engine (could be one of the above, or another engine you build and fit) render an opinion on whether residue from a smoldering crater indicates that the crater was caused by a fusion device if the instrument measures an alpha activity of 0.0061 counts per second.

## 2. Poison mushroom classification.

### Background

Mushroom foraging is a practice with both cultural significance and serious risks. Many wild mushrooms are edible and nutritious, but others contain potent toxins that can cause severe illness or death, even in small quantities. While advanced techniques like High-Performance Liquid Chromatography (HPLC) and mass spectrometry can accurately identify toxic compounds, these methods are laboratory-based, expensive, and time-consuming—making them impractical for real-time decision-making in the field. For amateur foragers, park rangers, and even emergency responders, the ability to assess toxicity risk based on visible features—such as cap color, gill shape, bruising reactions, and habitat—is far more practical, if not lifesaving.

This exercise seeks a machine learning classification algorithm to predict whether a mushroom is edible or poisonous based on easily observable characteristics. The training dataset is drawn from field guide-style descriptions, not lab tests, mimicking how a forager might gather information in the wild.

### Problem Statement

Using the database description at <http://54.243.252.9/ce-5319-webroot/2-Exercises/ES-2/agaricus-lepiota.names> and the encoded database <http://54.243.252.9/ce-5319-webroot/2-Exercises/ES-2/agaricus-lepiota.data> build a suitable prediction engine to classify a mushroom as poison or non-poison.

Apply your model on the mushroom *Coprinopsis atramentaria* which has the characteristics below<sup>2</sup>:

Measuring 3–10 centimetres in diameter, the greyish or brownish-grey cap is furrowed, initially bell-shaped, and later more convex, splitting at the margin. It melts from the outside in. The very crowded gills are free; they are white at first, then grey or pinkish and turn black and deliquesce.

The stipe measures 5–17 cm high by 1–2 cm thick, is grey in colour, and lacks a ring. In young groups, the stems may be obscured by the caps. The spore print is black and the almond-shaped spores measure 8–11 by 5–6  $\mu\text{m}$ . The flesh is thin and pale grey in colour.

It is commonly associated with buried wood and is found in grassland, meadows, disturbed ground, and open terrain from late spring to autumn. Like many ink caps, it grows in tufts.

---

<sup>2</sup>You will have to use my rather limited description to encode the mushroom, and determine its poison probability

Other information summaries include:

- Gills on hymenium
- Cap is ovate
- Hymenium is free
- Stipe is bare
- Spore print is blackish-brown
- Ecology is saprotrophic