

Midterm Exam Machine learning.

Last Name:

First Name:

1. Explain differences between supervised and unsupervised learning (20 p) (2 or 3 lines is enough)
2. Describe the relation between training error and overfitting/underfitting of the machine learning model. (20 p)
3. Considering the following observations:

| Day | Outlook  | Temperature | Humidity | Wind   | Play Tennis |
|-----|----------|-------------|----------|--------|-------------|
| D1  | Sunny    | Hot         | High     | Weak   | No          |
| D2  | Sunny    | Hot         | High     | Strong | No          |
| D3  | Overcast | Hot         | High     | Weak   | Yes         |
| D4  | Rain     | Mild        | High     | Weak   | Yes         |
| D5  | Rain     | Cool        | Normal   | Weak   | Yes         |
| D6  | Rain     | Cool        | Normal   | Strong | No          |
| D7  | Overcast | Cool        | Normal   | Strong | Yes         |
| D8  | Sunny    | Mild        | High     | Weak   | No          |
| D9  | Sunny    | Cool        | Normal   | Weak   | Yes         |
| D10 | Rain     | Mild        | Normal   | Weak   | Yes         |
| D11 | Sunny    | Mild        | Normal   | Strong | Yes         |
| D12 | Overcast | Mild        | High     | Strong | Yes         |
| D13 | Overcast | Hot         | Normal   | Weak   | Yes         |
| D14 | Rain     | Mild        | High     | Strong | No          |

---

Calculate the resulting class (Yes or No) that a Bayesian classifier would return for the following observation:  $x' = (\text{Outlook}=\text{Sunny}, \text{Temperature}=\text{Cool}, \text{Humidity}=\text{High}, \text{Wind}=\text{Weak})$

Build the learning phase and the test phase. (Help from supplemental material) (30 p).

4. Assuming that the RSS (a,b) get to be reduced to  $\text{RSS}(a,b) = a^2 + b^2 - 4a - 6b + 13$ . Calculate the best fitting linear model with the (gradient=0) method. (30 p).