

HW1 of Machine Learning with Python

Deadline: 15th of November (چهارشنبه ۲۴ آبان)

Email: mlfall2023@gmail.com



With the provided dataset, please complete the following homework assignments:

Six questions have been prepared for you to solve using this dataset. The description of the dataset is provided at the end of the questions.

1. Load the “**cellular network performance dataset**” into a pandas DataFrame and display the first 10 rows of the data.
 - a. Then, remove the extra indexing column
 - b. and, then, check for missing values and handle them appropriately column by column.
2. Calculate the summary statistics (mean, median, std, etc.) for the whole dataset. Then Visualize the distribution of 'Signal Strength' column using a histogram.
3. Filter the dataset to only include urban environment users and with less than 5 kms distance to tower.
4. Create a new column 'Call Duration (m)' representing the duration of the call in minute.
5. Group the data by the 'Environment' column and calculate the mean 'Call Duration (s)' for each group. Then, calculate the mean 'SNR' for each group.
6. Create a scatter plot showing the relationship between 'Signal Strength (dBm)' and 'Distance to Tower (km)'.

Dataset Features Information:

Features:

1. **Signal Strength (dBm):** Experience the power of wireless signals with precise measurements of signal strength, ranging from -50 dBm to -120 dBm. The signal strength indicates the intensity or amplitude of the received signal, allowing you to assess the quality and reliability of the network connection.
2. **Distance to Tower (kilometers):** Gain insights into the proximity of mobile devices to cellular towers. This feature provides an estimate of the distance between the device and the tower, enabling a deeper understanding of signal propagation and coverage patterns.
3. **SNR (Signal-to-Noise Ratio):** Delve into the signal quality with SNR measurements, representing the ratio between the signal power and noise power in the communication system. SNR values provide a quantitative assessment of signal clarity and robustness.
4. **Attenuation:** Explore the impact of signal loss as it traverses various mediums and encounters obstacles. Attenuation measurements shed light on the signal's weakening effect due to factors such as distance, interference, and environmental conditions.

Units:

1. **Signal Strength:** dBm (decibel milliwatts)
2. **Distance to Tower:** kilometers
3. **SNR:** unitless (ratio)
4. **Attenuation:** dB (decibels)