Systems Neuroscience 1st Assignment – Neural coding in the visual system
Submission Deadline: 11/02/2024

General guidelines:
Please submit on moodle all your answers in a pdf format, including all the required figures for each question, along with the code.
Make sure your plots include titles, axes with units and legends and you answer all question parts.

The following exercise contains data from neuronal activity recordings of larval zebrafish using calcium imaging. In this experiment, neuronal activity of the optic tectum was recorded while the fish was presented with prey-like stimuli, in the form of stationary and moving dots, in varying azimuthal locations in its field of view. Each stationary stimulus was presented for 1.5 seconds. The optic tectum is the main visual processing area of the zebrafish, analogous to the superior colliculus of the mammalian brain.

The data contains the following fields:

- **dff**: The calcium signal of each neuron. A 2d array, for which the first axis is the different neurons (1089 neuron) and the second axis is the different frames in the experiment (17,000 frames with sampling rate of 3 Hz)
- **centers**: The location of the centers of the neurons of the zebrafish brain, in pixels. Includes a 2d array, for which the first axis is the different neurons, and the second axis is the location of each neuron (first the x pixel, second the y pixel).
- **stimuli**: The azimuthal location (in degrees) and the onset frame of each of the presented stimuli during the experiment. Includes a 2d array, for which the first axis is the different stimuli (215 stimuli) and the second axis is the information about each stimuli (first the azimuthal location, second the onset frame).

Questions:

1. Data exploration - visualize the signals and their relation to the stimuli
   a. Plot calcium signals for 5 exemplary neurons separately between the 5000-th frame to 8000-th frame. Add stimuli onset and duration as a colored rectangle to the previous plot (hint: ax.axvspan with alpha parameter). What can you deduce about the neuronal activity from this plot?
   b. Plot the population response (i.e., calcium signal from all neurons, hint: matshow). To do so, first cap neuronal activity at 5. Add a color bar to your plot. What information do you gain by looking at the population response?
   c. Add stimuli onset and duration as a colored rectangle to the previous plot, focusing on the neuronal activity between the 5000-th and 6000-th frame. How do you interpret the results with this addition?
   d. For each neuron (i) for each stimulus (j), calculate the mean neuronal activity (of the raw data), by calculating the average calcium response across 9 frames from the

stimulus onset (a matrix A of neurons X stimuli, e.g. 1089 X 215). Include only stationary stimuli in the next sections. This matrix will be used for the next sections.

2. <u>Tuning curves</u>

   a. Plot the tuning curve (i.e., the average neuronal response across trials for each of the different stimuli from 1d), of 5 exemplary neurons, including SDs. How variable are the tuning curves between neurons? What is the preferred stimulus of each neuron based on this plot? What can you deduce from that? Please note that some neurons may not have a strong preference.
   b. Calculate and plot the fraction of neurons for each preferred stimulus type. Can you deduce the preferred stimulus or stimuli of the zebrafish according to this analysis? Try to hypothesize why it is the case.
   c. Based on the data regarding the locations of neurons centers of the zebrafish brain, plot neurons locations on an xy plane, colored by their preferred stimulus. What new information do you gain from this plot? Try to hypothesize why it is the case.

3. <u>Correlation in the data and dimensionality reduction</u>

   a. Calculate the z-score for each neuron (based on matrix A in 1d). Run dimensionality reduction using PCA with 2 components, project the data on the first 2 PCs, and plot your results (hint: sklearn). Show the percentage of variance explained by each component. What is the meaning of the two PCs? What do each dot in the plot represent? What do we learn from such dimensionality reduction analysis? Color each dot according to the preferred stimulus, with a color scale indicative of the azimuthal location. What can you deduce from this result?
   b. Transpose the average neuronal activity (A') and repeat the previous PCA reduction analysis. Color each dot according to stimuli, and plot the dimensionality reduction with an appropriate color bar. Make sure your color palette is indicative of the stimulus azimuthal locations. What is the meaning of the two PCs? What do each dot in the plot represent? What information do we gain from this plot compared to the previous one?
   c. Another way to visualize data in the lower dimensional space is using tSNE. Run tSNE with 2 components over the transposed neuronal activity (A'). Color each dot according to the respective presented stimulus, and plot the results with a color bar indicative of the azimuthal location. What does each dot represent? How do these results differ from or similar to the previous ones? Hypothesize what conditions are suitable for tSNE and what for PCA, why?
   d. What can you say regarding information stored in single neurons compared to several neurons based on this analysis?

<u>4. Decoding the presented stimulus using neural activity only</u>

   a. From matrix A select only neurons that were selective to stimuli 75 and -75 degrees. Fit an LDA model to the PC space of the activity of these selected neurons. Repeat the PCA plot, with a boundary line based on the results of the LDA algorithm (hint: mglearn). How well does the model discriminate between the two stimuli? How can you measure this?

b. Plot a confusion matrix of the predicted labels (i.e., two stimuli) of the data included in 4a.
c. What is the accuracy and sensitivity?
d. Repeat the same analysis for the two stimuli -75 and -60 and plot your results. How well does the model discriminate between the two stimuli in this case? How good is the LDA algorithm in performing this discrimination task in general? why?
e. What assumptions do you have in order to use LDA for this data? Are they met?
f. Suggest another analytical approach to find the preferred stimulus or stimuli of the fish, not mentioned in this exercise. Explain briefly how you will do it.