LIGN 167 final projectModel architecture

Date Created	@November 27, 2022 2:03 PM
Status	

Model architecture

- Option 1
    - Use model_data.csv.
    - Encoder - decoder transformers. (translation from photometry to analog)
        - Training: seq of gcamp_lp and seq lp_duration (or lp success or not).
        - Testing: given a seq of gcamp_lp, what is the expected seq of lp durations.
    - Good reasons:
        - Data in a nice temporal sequence → easier to train and test.
    - Bad reasons:
        - Uninteresting because **maybe** (or maybe not) that a great gcamp_lp means successful lp.
        - gcamp_lp and lp_duration are continuous data.
- Option 2
    - Use raw gcamp and analog data.
    - Encoder - decoder transformers. (translation from photometry to analog)
        - Training: seq of gcamp activity and binary seq (lp or not).
        - Test: given a seq of gcamp activity, whats the the expected binary sequence.
    - Good reasons:

        - Granular information of the data. More rows = better model? and possible less obvious relation between gcamp and lp.
    - Bad reasons:
        - Not a straight forward way to match both data sources because of different timestamps. Relation between the two data sources may vary across mice.
            - Possible solution: subset data when the two sources have records.
        - gcamp is continuous data.
        - Given the model prediction on sequence of lever presses, what is the timeframe in which those prediction occur?
            - Possible solution: assume each predicted event happens 4 milliseconds (eliminating decimals, each step is 4 milliseconds apart the next one) after the previous event.
- Option 3
    - Use encoder - decoder transformers in a timeseries application.
        - Encoder is a sequence of n lever presses.
        - Decoder is the first k lever presses of the encoder sequence.
    - Bad reasons:
        - How can we incorporate the photometry data?

Issue

- Distribution in between sessions do not seem similar