Nathaniel Nyema
CNS187
Final Project Progress Report
25 May, 2023

The topic I chose for my final project was topic 2: Models of recognition memory with single trial learning. The goal of this project is to perform a detailed study of the paper Meta-learning synaptic plasticity and memory addressing for continual familiarity detection by Tyulmankov et al. and test how well the proposed network can generate. So far I have read through this paper and was able to successfully reproduce some of the results. Specifically, I first wrote a function in Python to generate inputs to the network as described in the paper (Fig 1) such that I could reproduce some of the findings the authors presented. This task is essentially a recognition memory task where the network is sequentially presented with a random binary string with probability 1-p. With probability p it is instead presented with the input it saw R time steps prior if that input has not been repeated as yet. However, if it this putative input has already been repeated a random string is generated. Generally p was set to 0.5 for all experiments.

```python
generate_input = lambda d: 2*np.round(np.random.rand(d)) - 1

def generate_dataset(T, R, d, p):
    x = np.zeros((T,d))
    y = np.zeros(T)
    for i in range(T):
        repeat = False
        if (i>=R) and (np.random.rand()<p):
            if not y[i-R]:
                repeat=True
        if repeat:
            x[i] = x[i-R]
            y[i] = 1
        else:
            # generate random input but make sure
            # we haven't seen it yet
            # to be fair it's very unlikely we will generate
            # a random string we've already seen (p = i/(2^d))
            # but just in case)

            if i>0:
                in_x = True
                while in_x:
                    _x = generate_input(d)
                    in_x = (_x[None,:] == x[:i]).all(axis=1).any()
            else:
                _x = generate_input(d)
            x[i] = _x
    return x, y
```

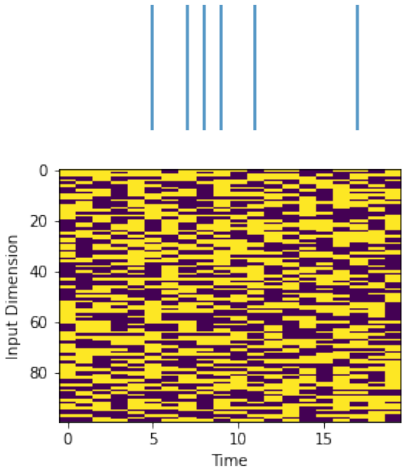

Figure 1 The recognition memory task. (left) code used to generate the dataset. (right) example dataset with R=5 and p=0.5.

I then built the idealized HebbFF network following the prescriptions towards the end of the paper which require no meta-learning. Specifically, this involved defining a static weight matrix W1, a plastic one A, the plasticity rules and a bias value. W1 functions as an addressing matrix for different combinations of a subset n of the d inputs to the network, such that each hidden unit receives a unique combination of these scaled binary inputs. The paper goes into further details on how this scaling factor is chosen. All the weights in A are set to 0 at the start of a given task and evolve according to the equation $A(t+1) = \lambda A(t) + \eta h(t)x(t)^T$, where η is set to -1 in the idealized model λ is set to according to equations detailed in the paper. The bias value is chosen such that only one unit in the hidden layer is active at once. Given these hyperparameters, the output of the hidden units is defined by the following equation:

$$h(t) = \Theta((\mathbf{W}_1 + \mathbf{A}(t))x(t) + b_1)$$

Where Θ is the heaviside function. We make the prediction that the network has seen a given string if all hidden units have output 0. Otherwise, we say that we have not seen the string. The specific network I used took in a 100-dimensional input, 5 of which were designated as

addressing input, the rest of which had plastig synapses onto the 2^5 = 32 hidden units. This network was subjected to the aforementioned task for different values of R and for different dataset lengths (T) to reproduce figure 3C,D (Figure 2). I also show an example of the hidden unit activities over time in Figure 3.

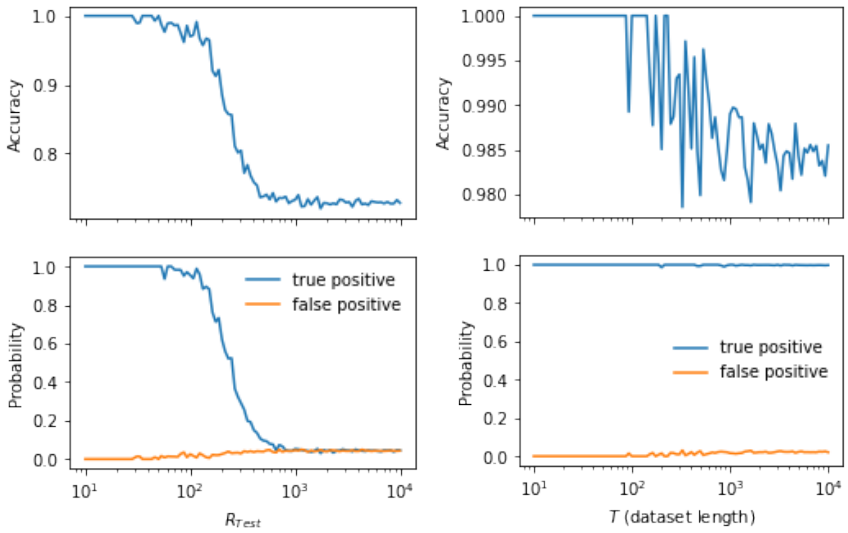

Figure 2 Reproducing Figure 3c-d from the paper. On the left is the networks performance as a function of the interval R in the recognition memory task. Dataset length is set adaptively here to 3*R. On the right are the same stats for R=5 at different dataset lengths

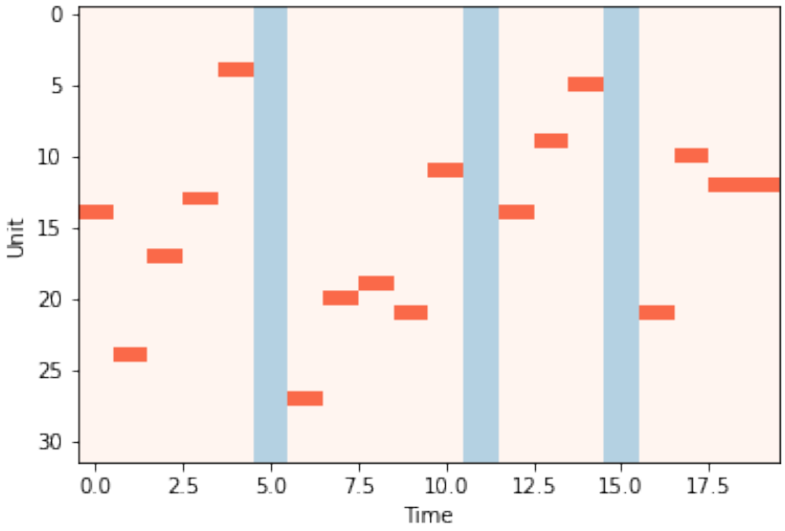

Figure 3 Reproducing 4a. The activity of the hidden units is shown in red here as a function of time. Blue regions highlight when the trial was a repeat trial.