RNN Example: Predicting Next Day Food Sequence

Introduction

This document illustrates a simple Recurrent Neural Network (RNN) model that predicts the sequence of next day's food consumption based on the previous day's food choices. Let:

- x_t be the food vector at day t
- h_t be the hidden state at day t
- y_t be the predicted food vector for the next day $t+1$

Mathematical Model

The RNN updates are given by:

$$h_t = \tanh(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \tag{1}$$
$$y_t = W_{hy}h_t + b_y \tag{2}$$

where:

- W_{xh}: weight matrix from input to hidden layer
- W_{hh}: weight matrix from hidden to hidden layer
- W_{hy}: weight matrix from hidden to output layer
- b_h, b_y: bias vectors

Matrix Representation

Assume a simple case with 3 types of food (e.g., *Chicken, Pizza, Pasta*) represented by one-hot encoded vectors:

$$x_t^{(1)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{(Chicken on day } t\text{)} \quad x_t^{(2)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{(Pizza on day } t\text{)} \quad x_t^{(3)} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{(Pasta on day } t\text{)} \tag{3}$$

The hidden state h_t is updated as:

$$h_t = \tanh\left(\begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} x_t + \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{bmatrix} h_{t-1} + b_h\right) \tag{4}$$

Finally, the output y_t predicts the next day's food:

$$y_t = W_{hy} h_t + b_y \tag{5}$$

If the model predicts *Burger* for the next day, the output might be:

$$y_t = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ (Pizza on day } t+1\text{)} \tag{6}$$

Loss Function

We use cross-entropy loss to measure the distance between the predicted food vector y_t and the actual food vector y_t^{true}:

$$\text{Loss} = -\sum_i y_t^{true}[i] \log(y_t[i]) \tag{7}$$

where y_t^{true} might be the correct next-day food, e.g., Pizza:

$$y_t^{true} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \tag{8}$$

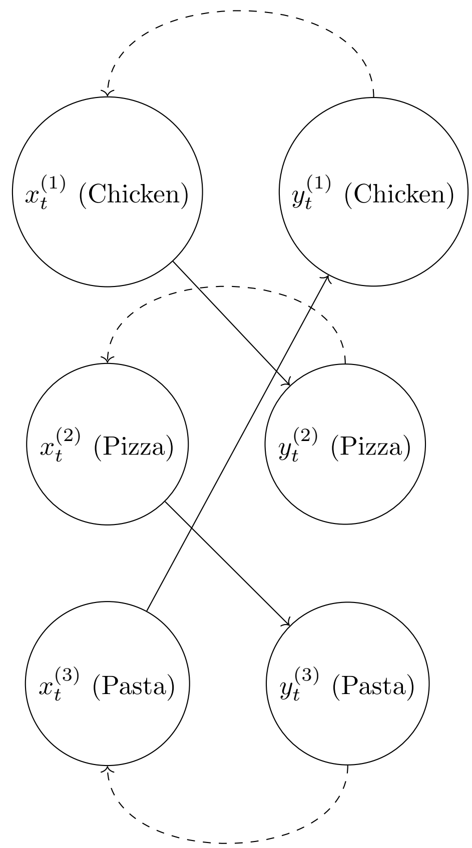
Connecting Previous and Next Items

The relationship between the previous and next items can be visualized as a sequence of interconnected nodes, where the hidden state h_t carries information from the previous time step $t-1$ to the next step $t+1$ through recursive connections. This allows the RNN to learn dependencies over time.

The full sequence can be expressed as:

$$h_t = \tanh(W_{xh} x_t + W_{hh} h_{t-1} + b_h) \tag{9}$$
$$h_{t+1} = \tanh(W_{xh} x_{t+1} + W_{hh} h_t + b_h) \tag{10}$$
$$y_t = W_{hy} h_t + b_y \tag{11}$$
$$y_{t+1} = W_{hy} h_{t+1} + b_y \tag{12}$$

Visualization



The Story

Once upon a time, there was a student who lived near a small shop run by a friendly shopkeeper. The shopkeeper sold delicious chicken, burger and pizza. Each day, the student would go to the shop — but only if the weather was sunny. On rainy days, the student stayed home and ate whatever food he had from the previous day.

Every day, the shopkeeper would ask the student:

> *"What you had on the previous day? Try the new item on my list."*

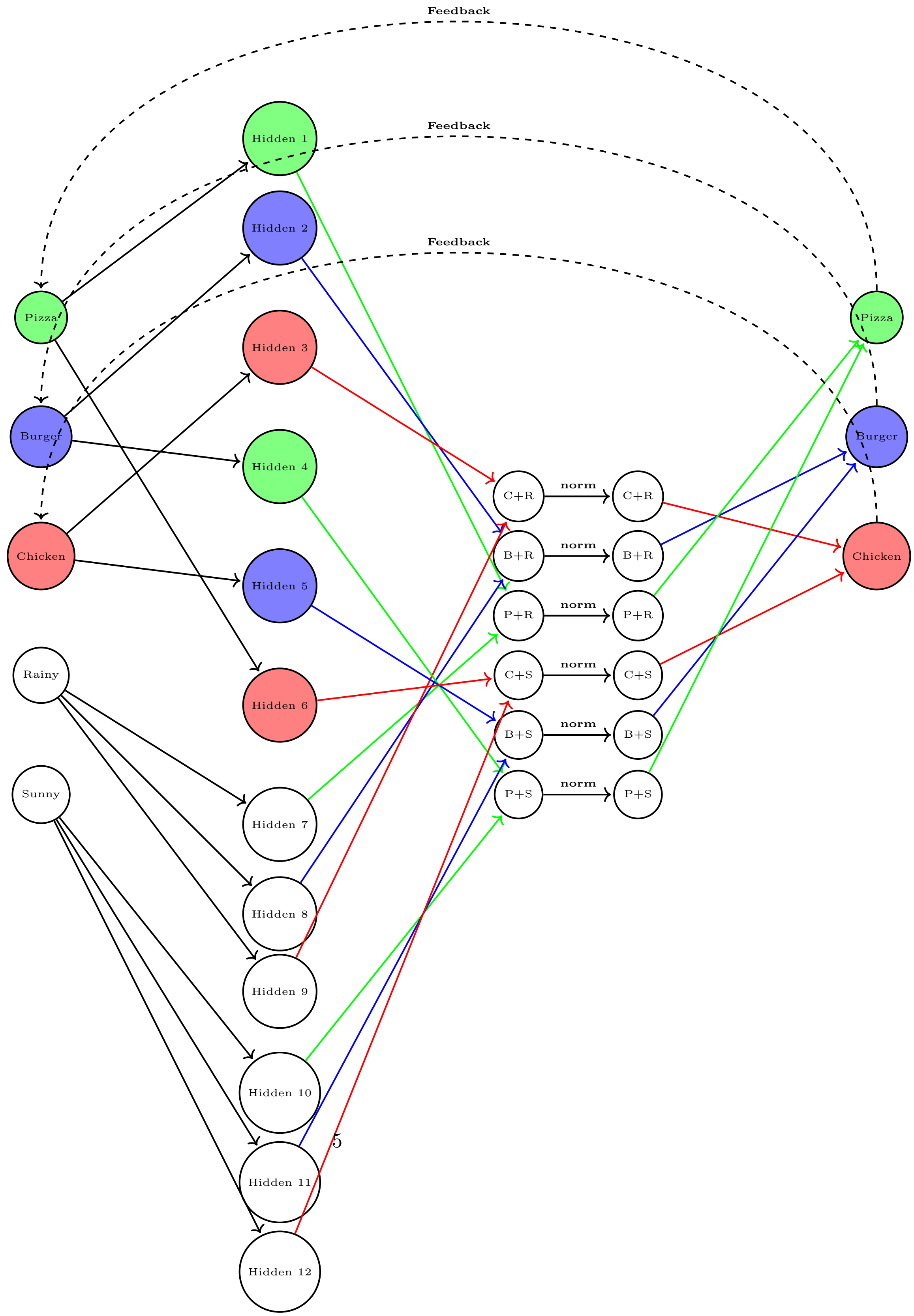
The student would answer what he had, but if it rained, he had to stick with what he ate the day before.

The Weekly Plan

Below is a record of what the student ate each day, considering both the weather and his choices.

Day	Weather	Prev Day Food	Next Day Food
Monday	Sunny	Chicken	Burger
Tuesday	Rainy	Burger	Burger
Wednesday	Sunny	Burger	Pizza
Thursday	Rainy	Pizza	Pizza
Friday	Sunny	Pizza	Chicken
Saturday	Rainy	Chicken	Chicken
Sunday	Sunny	Chicken	??

The student enjoyed his meals and learned that sometimes, the weather could decide his menu just as much as his own cravings.


Feedback
Feedback
Feedback
Pizza
Burger
Chicken
Rainy
Sunny
Hidden 1
Hidden 2
Hidden 3
Hidden 4
Hidden 5
Hidden 6
Hidden 7
Hidden 8
Hidden 9
Hidden 10
Hidden 11
Hidden 12
C+R
B+R
P+R
C+S
B+S
P+S
norm
norm
norm
norm
norm
norm
C+R
B+R
P+R
C+S
B+S
P+S
Pizza
Burger
Chicken