Prediction of Extreme Climate Conditions in Namal Valley using Machine Learning

Malik Jahan Khan, PhD
Department of Computer Science
Syed Babar Ali School of Science and Engineering
Lahore University of Management Sciences

Background

- Namal lake is vulnerable to flood in the monsoon seasons, in other parts of the year it faces water scarcity.
- These challenges resulted in fatalities, damage to the infrastructure and low agricultural yield over years.
- The two very recently flooding events occurred in 2015 and 2020 respectively. In 2022, the dam was facing severe drought.

DAWN

)6, 2023

Pipes being used to siphon water from Namal Lake

The Newspaper's Correspondent | Published August 23, 2015

0

MIANWALI: The Namal Lake flooding has inundated several localities since Aug 6 after the lake started overflowing from all sides as its gates are jammed.

ill torrent victims await relief

ugust 8, 2015

0

ter is overflowing the dam. — Dawn

I: Nothing is being done to ease the plight of residents of the Namal have been affected by hill torrents .

Early Warning System

- Water Level Prediction
- Drought Prediction
- Calculating Risks
- Alert Communication
- In-time Safety Actions

- Risk Knowledge
- Monitoring/ In-situ Sensors
- Meteorological Dataset
- Deep Learning Techniques
- Hydrological Models

13 CLIMATE ACTION

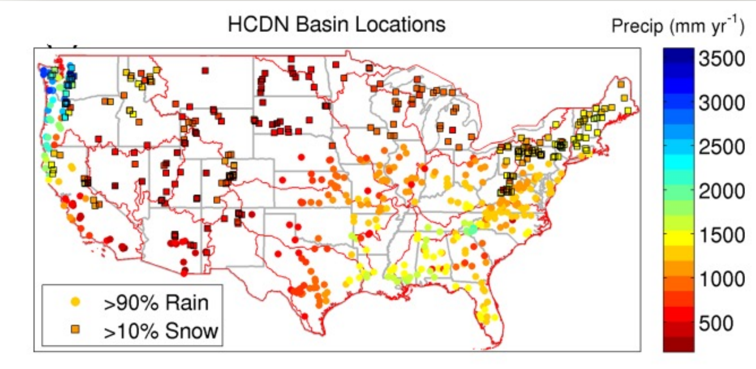
11 SUSTAINABLE CITIES AND COMMUNITIES

Methodology

1. Collecting Dataset From Different Sources
2. Pre-processing The Dataset
3. Building The Deep Learning Models
4. Testing and Model Optimization
5. Using Results For Risk Assessment
6. Alert Communication Services

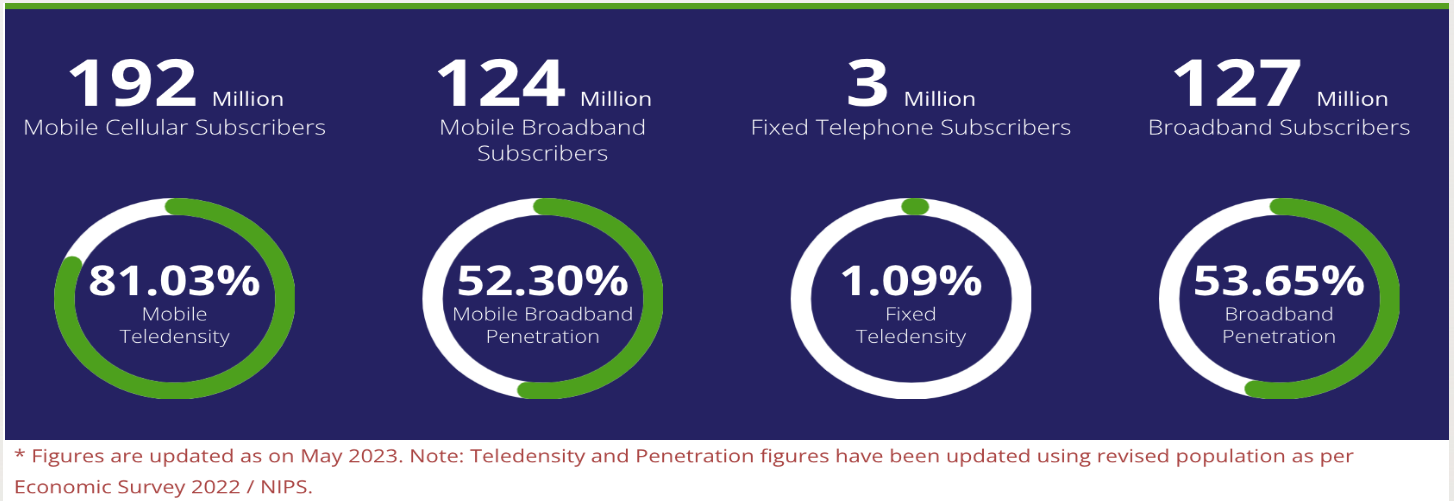
Data Sources

- Data from sensors installed by **Centre For Water Informatics And Technology, LUMS.**
- 5 years of historical data from irrigation office.
- Catchment attributes and meteorological for large sample studies (CAMEL) data set.
- 37+ years of rainfall dataset from **Center Of Hydrometeorological And Remote Sensing** data portal.



The Opportunity

- Smart phone penetration in the community
- Impact of "Machine Learning" in our daily life applications

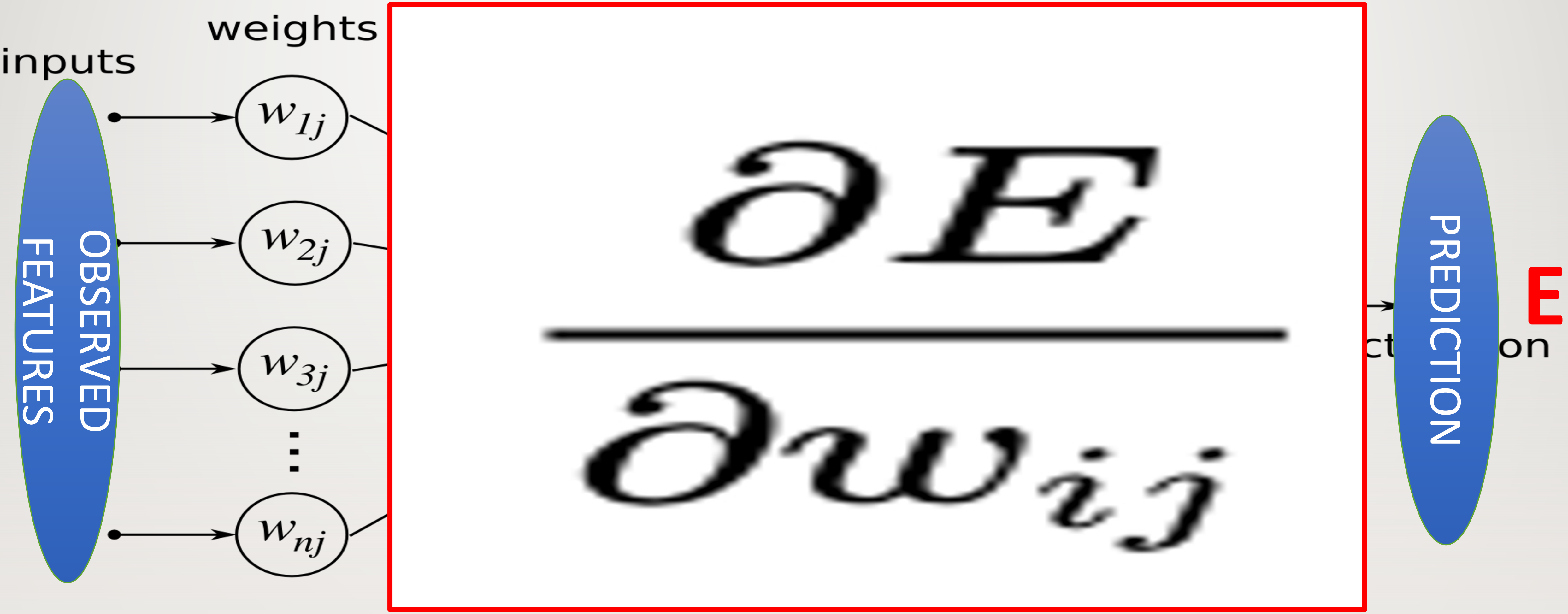


Proposed Solution

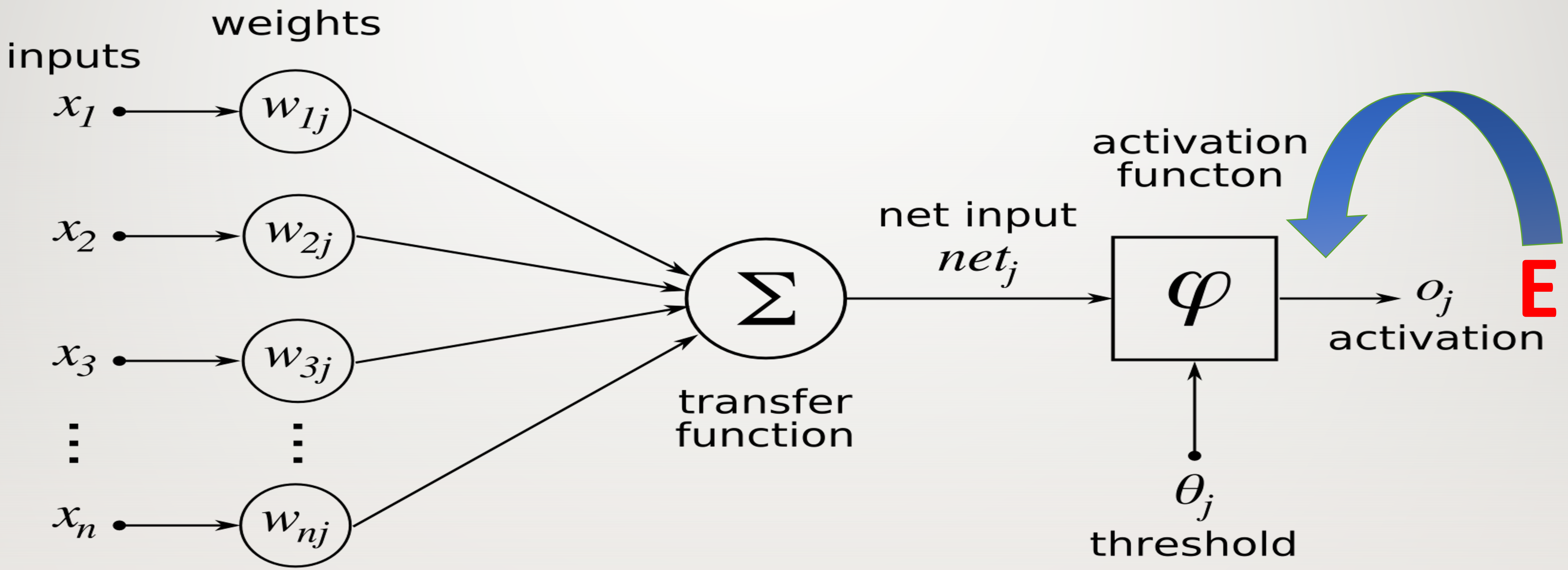
- Leverage neural networks / deep learning
- Train the machine learning models on the data collected from different sources
- Forecast the extreme climate conditions in the valley
  - Flash flood forecast
  - Drought forecast

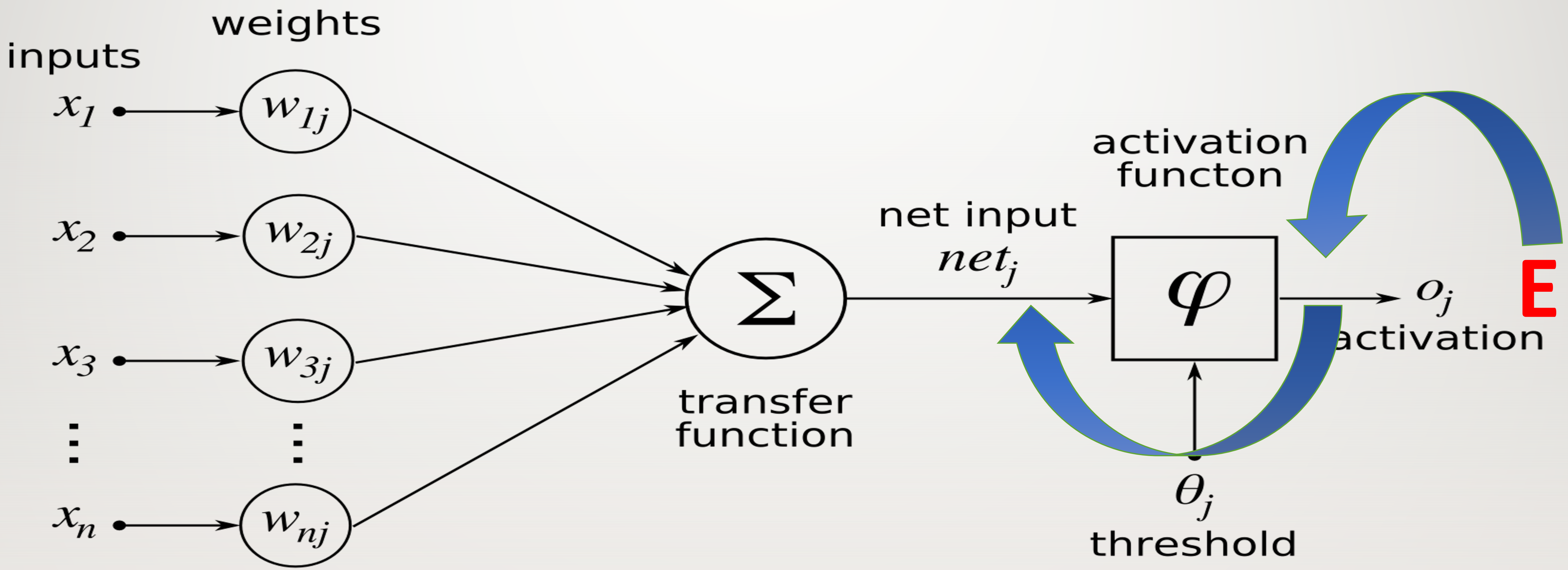
Implementation

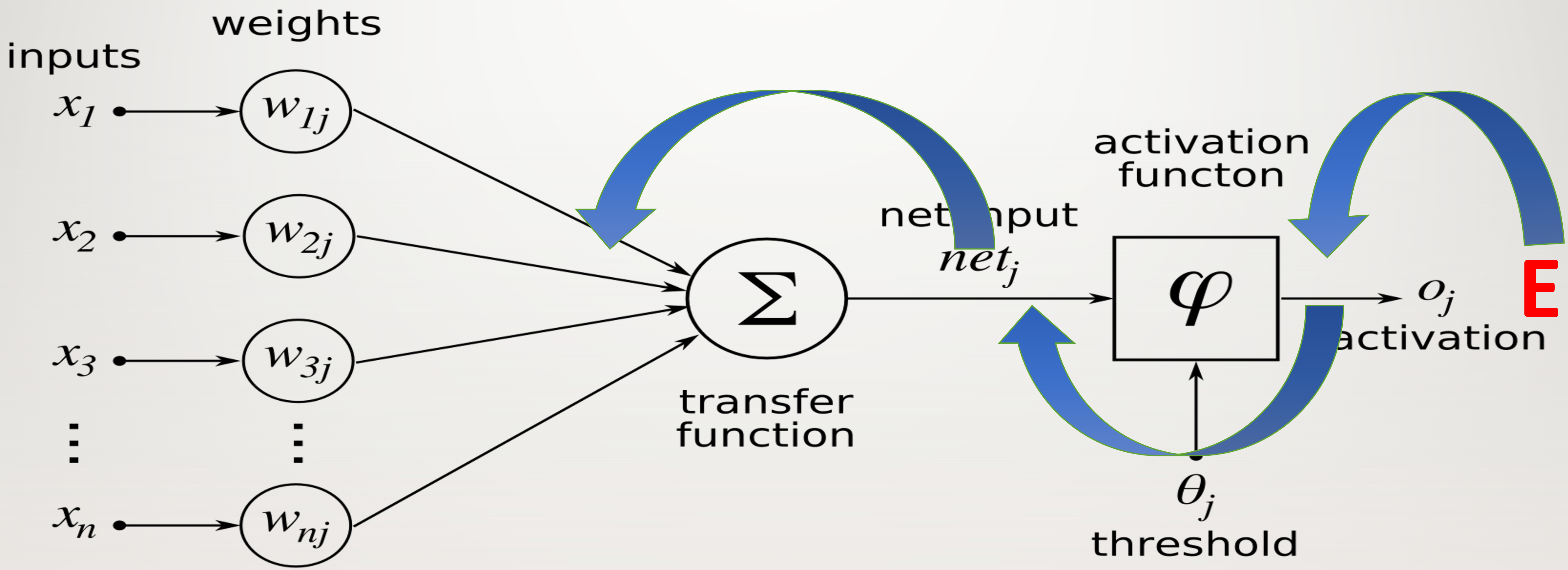
- Python
- TensorFlow / CNN / RNN / LSTM
- Amazon SageMaker Studio
- Firebase Database
- Android Studio



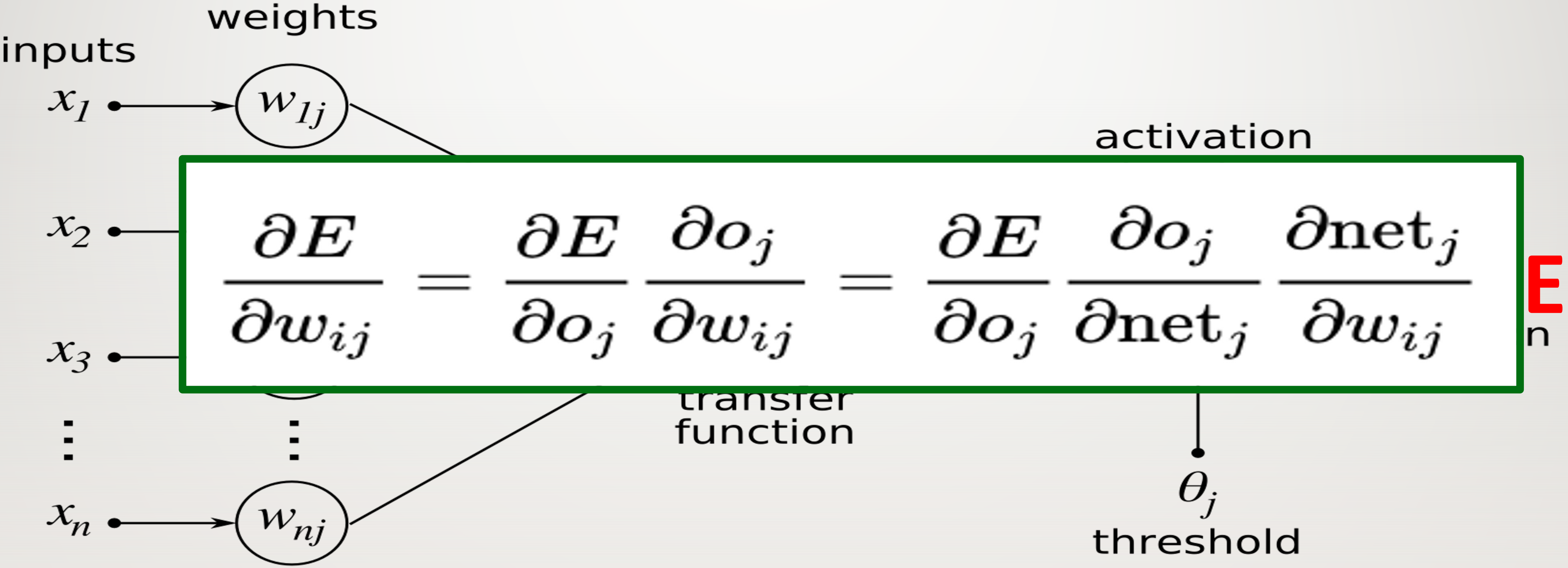
$$\frac{\partial E}{\partial w_{ij}}$$


weights
inputs
x_1
w_{1j}
x_2
w_{2j}
x_3
w_{3j}
x_n
w_{nj}
Σ
transfer function
net input
net_j
activation functon
φ
o_j
activation
E
θ_j
threshold


weights
inputs
x_1
w_{1j}
x_2
w_{2j}
x_3
w_{3j}
x_n
w_{nj}
Σ
transfer function
net input
net_j
activation functon
φ
o_j
activation
E
θ_j
threshold

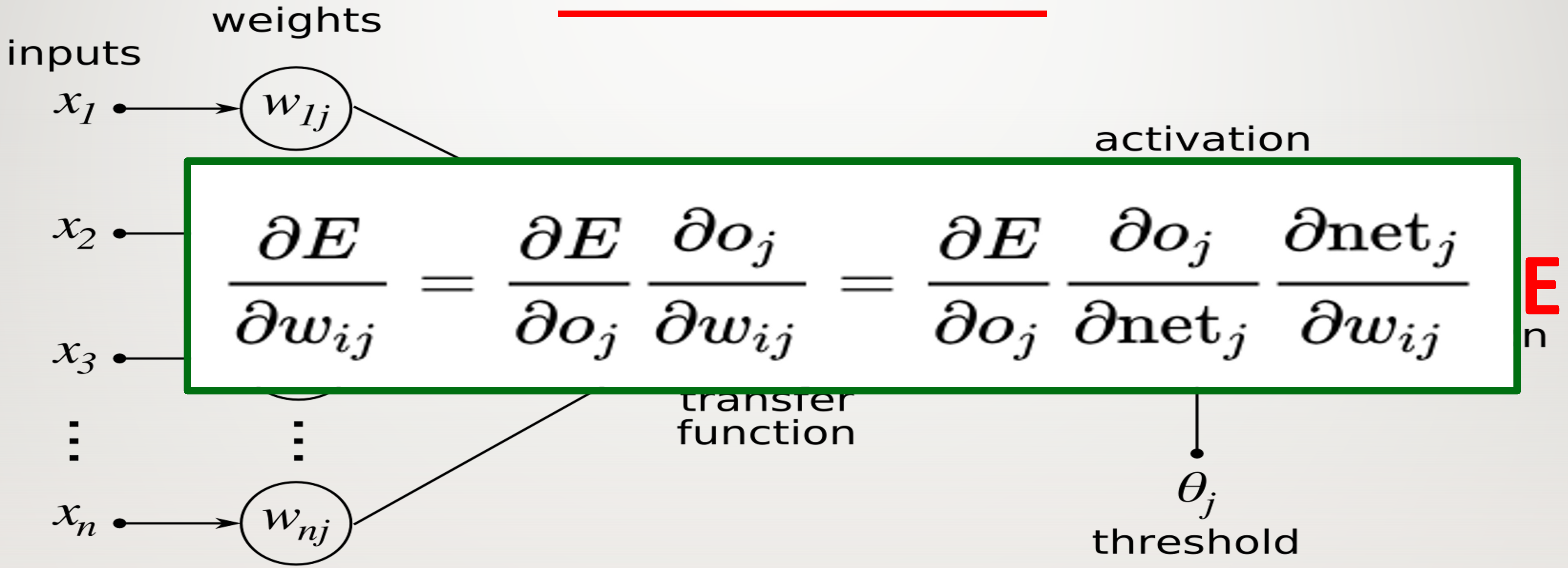


weights
inputs
x_1
w_{1j}
x_2
w_{2j}
x_3
w_{3j}
x_n
w_{nj}
Σ
transfer function
net input
net_j
activation functon
φ
o_j
activation
E
θ_j
threshold



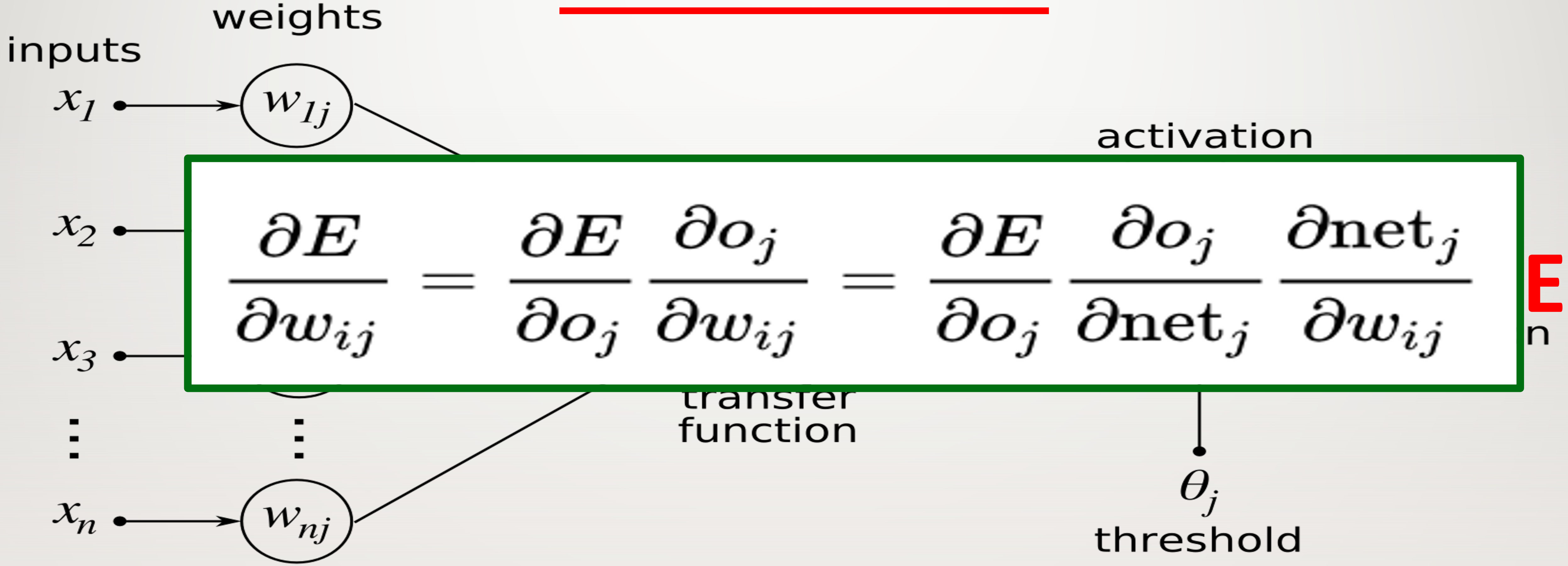
$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial o_j}\frac{\partial o_j}{\partial w_{ij}} = \frac{\partial E}{\partial o_j}\frac{\partial o_j}{\partial \mathrm{net}_j}\frac{\partial \mathrm{net}_j}{\partial w_{ij}}$$

Chain Rule



$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial o_j}\frac{\partial o_j}{\partial w_{ij}} = \frac{\partial E}{\partial o_j}\frac{\partial o_j}{\partial \mathrm{net}_j}\frac{\partial \mathrm{net}_j}{\partial w_{ij}}$$

Chain Rule



$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial o_j}\frac{\partial o_j}{\partial w_{ij}} = \frac{\partial E}{\partial o_j}\frac{\partial o_j}{\partial \mathrm{net}_j}\frac{\partial \mathrm{net}_j}{\partial w_{ij}}$$

Rate of Change of Error with respect to w_{ij}

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}}$$

Apply update rule!

$$w'_{ij} = w_{ij} + \Delta w_{ij}$$

Back-propagation Algorithm: Artificial Neural Networks

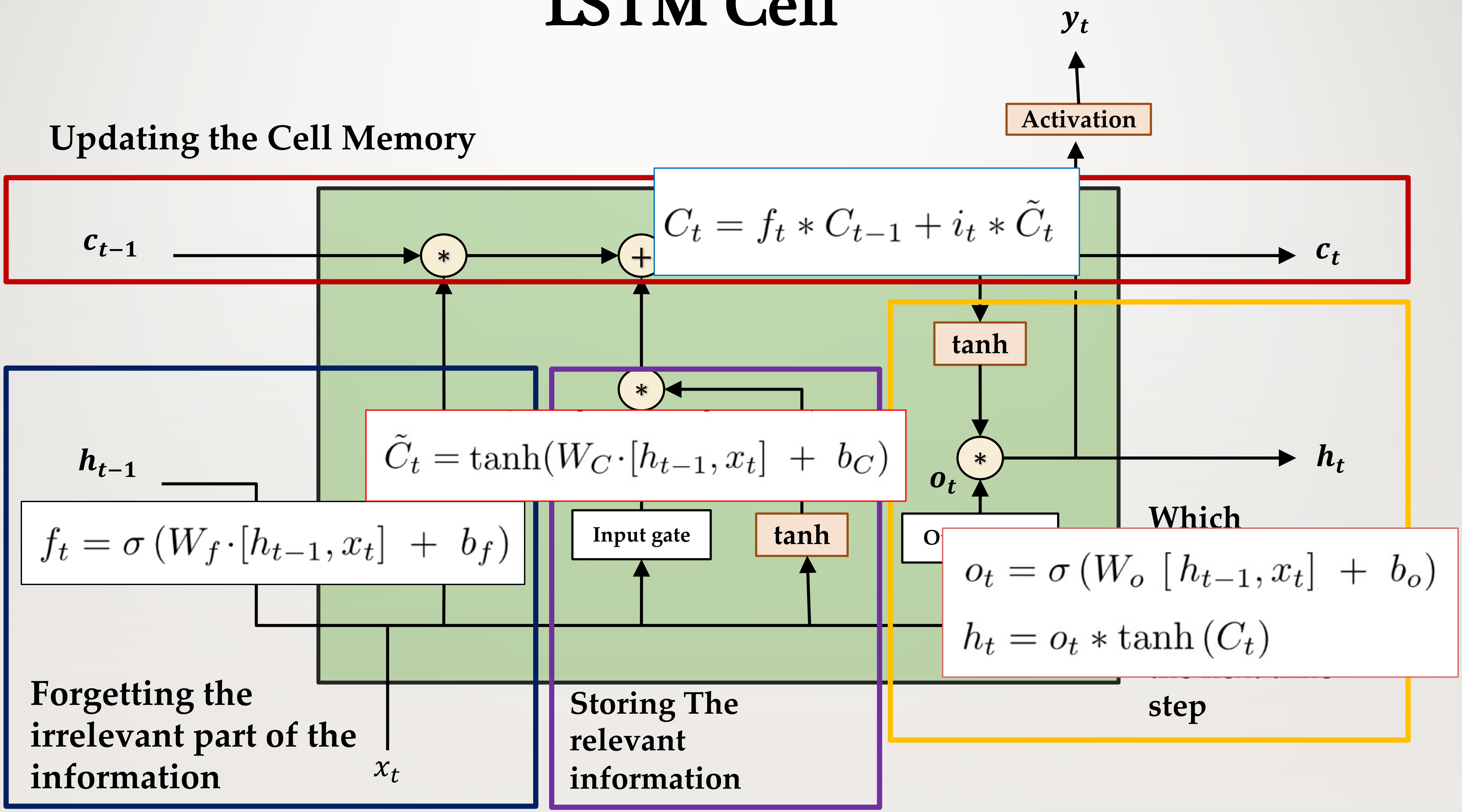
Recurrent Neural Networks (RNNs)

- Conventional neural network takes a bunch of input features and gets itself trained to estimate the output based on the training data

- Bottleneck: It takes the input features all at once and doesn't discriminate the features based on the temporal history

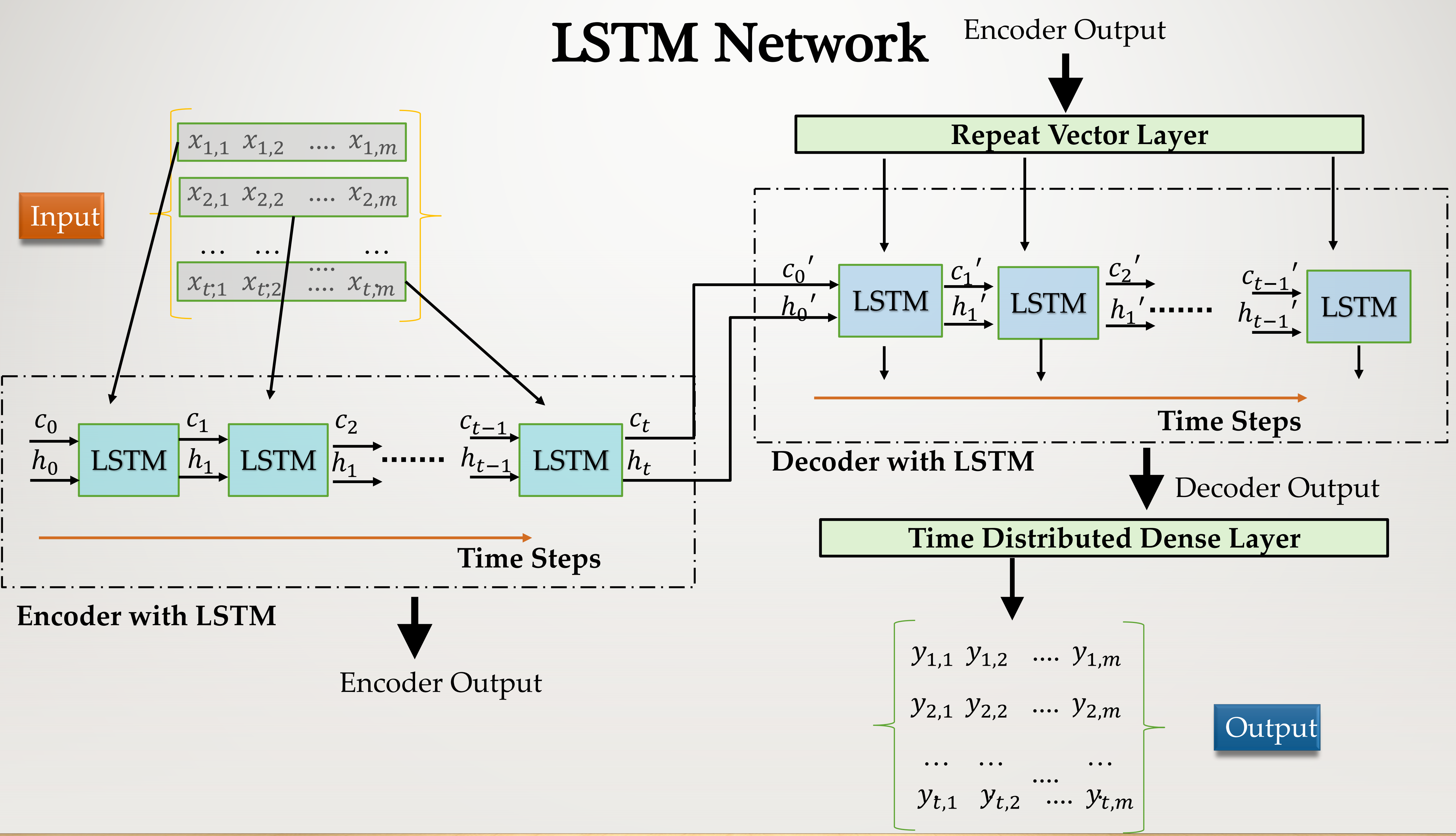
- RNNs are used to resolve this bottleneck!
  - **Long Short-Term Memory (LSTM)** is one of the popular RNNs

LSTM Cell



Updating the Cell Memory
c_{t-1}
$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$
c_t
y_t
Activation
tanh
h_{t-1}
$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$
$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right)$
Input gate
tanh
o_t
h_t
Which
$o_t = \sigma\left(W_o \left[h_{t-1}, x_t\right] + b_o\right)$
$h_t = o_t * \tanh\left(C_t\right)$
step
Forgetting the irrelevant part of the information
x_t
Storing The relevant information

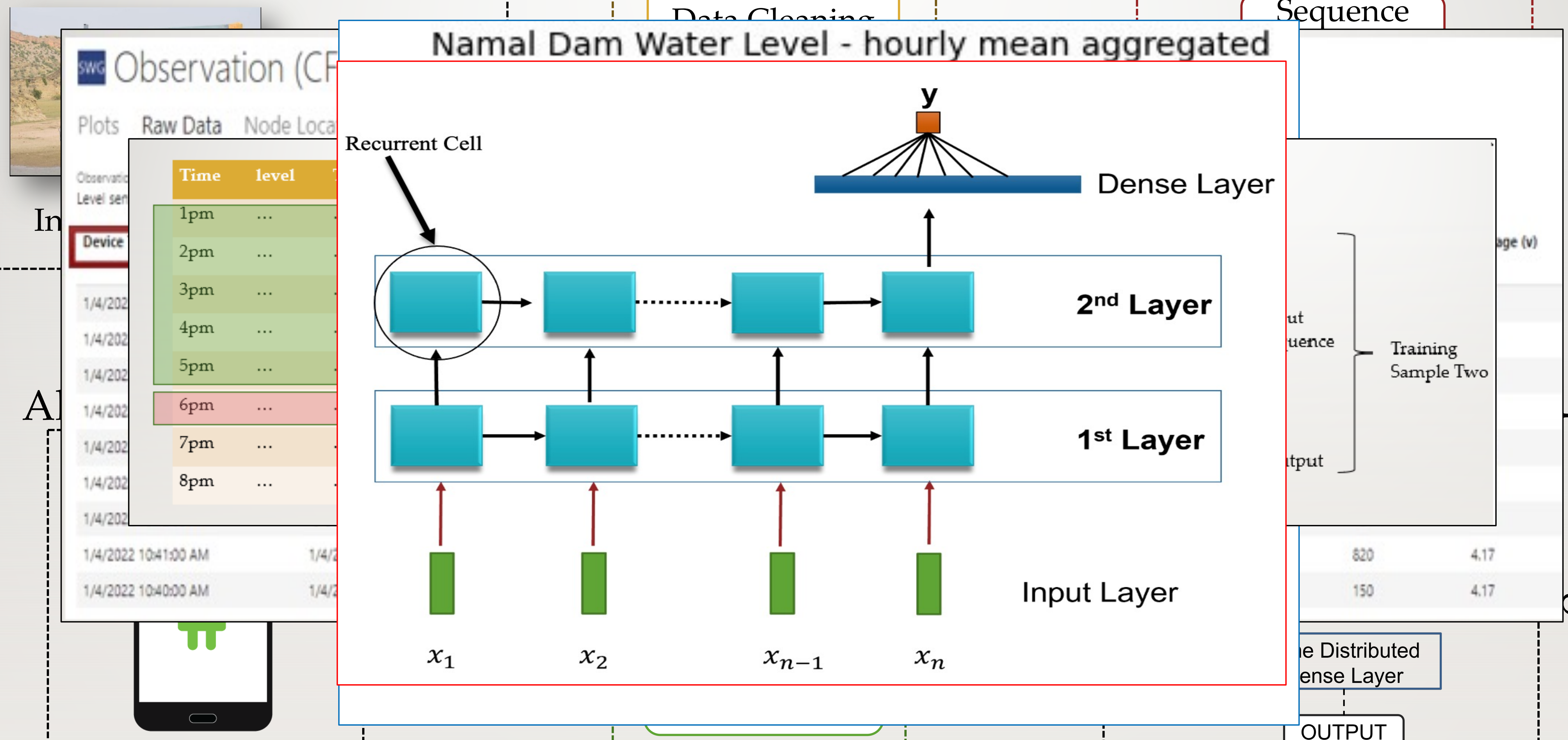
LSTM Network



Monitoring and Data

Pre-Processing Block

Data Pre-settings



Mobile Application

Deep Learning Model

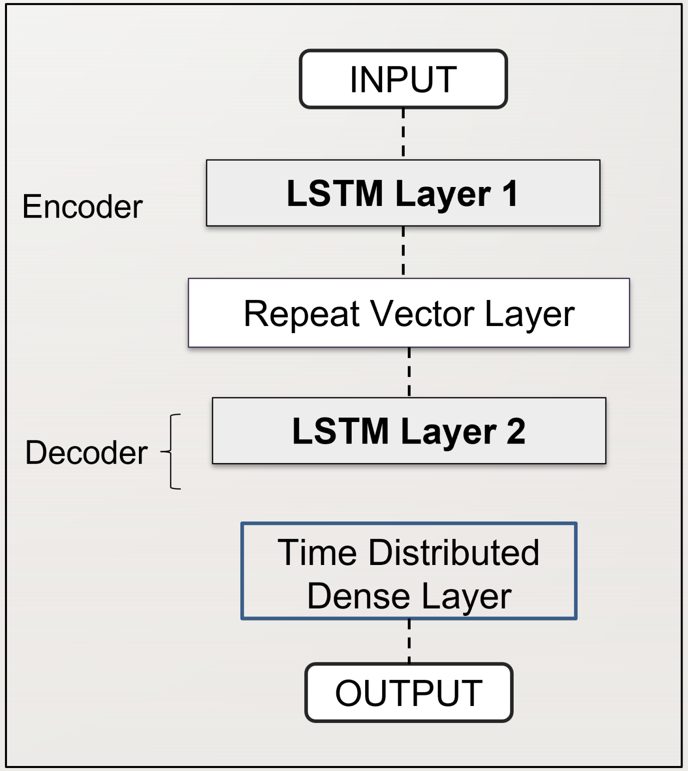
Water Level Prediction Model

Training Process	
Number Of LSTM Layers	2
Hidden States In layers	10
Activation	Tanh (tangent hyperbolic)
Batch Size	10
Input	144 hours data
Number Of Features	4
Number Of Epochs	20
Optimizer	Adam
Loss Function	Mean Square Error
Output	Next 5 Hours data points

5-Hours (short term flood forecasting)

Features

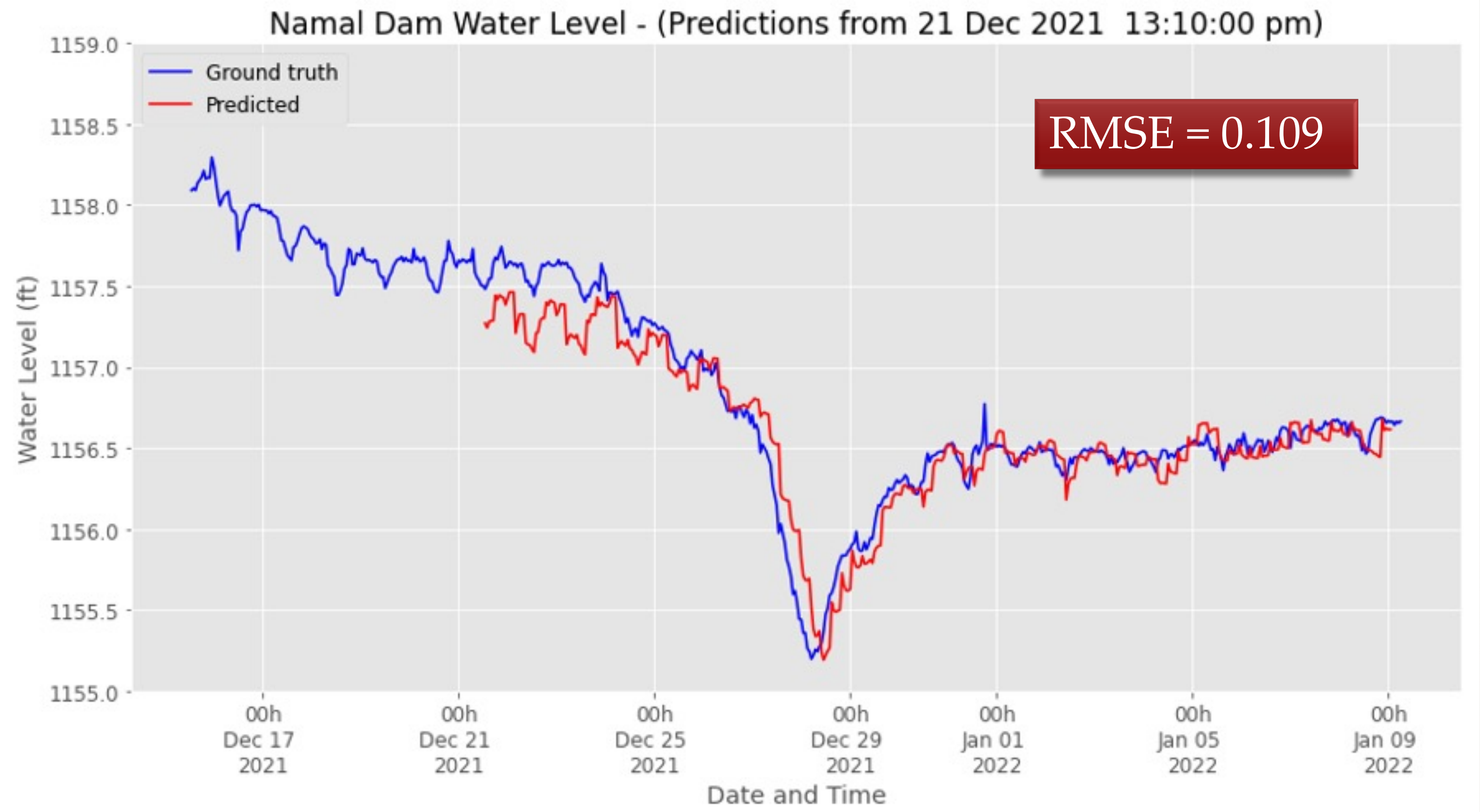
- ❑ Lake Water Level
- ❑ Rain
- ❑ Temperature
- ❑ Humidity



Deep Learning Model

Results

LSTM input = 144-hours (6-days) , Prediction = Next 5 hours



Results Comparison

Lake Water Level Prediction Model	RMSE on different input size		
	144h (6 days)	72h (3 days)	24h (1 day)
Hourly max	**0.109**	**0.114**	**0.119**
Hourly average	**0.185**	**0.191**	**0.182**

Risk Assessment

AREA UNDER LAKE	
Lake Water Level (ft)	Area In Acres
1160	1338.39
1165	1942.6
1170	2830.60
1175	3721.33

Flood Risk Classes	
Lake Water Level (ft)	Risk
Less than 1162	Normal
1162.1 – 1165	Low
1165.1 – 1168	Medium
1168.1 – 1170	High
Above 1170	Extreme

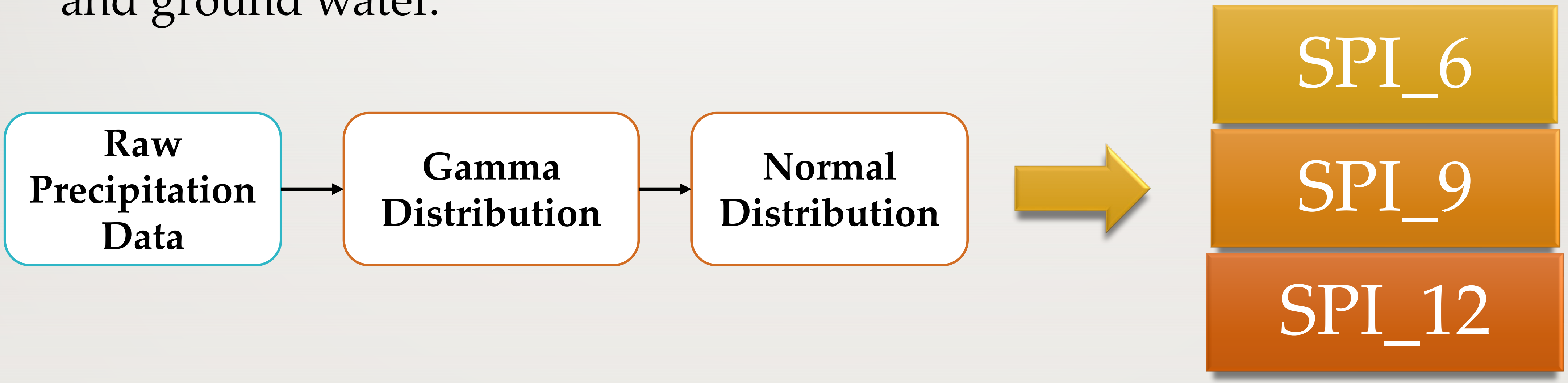
Irrigation Department

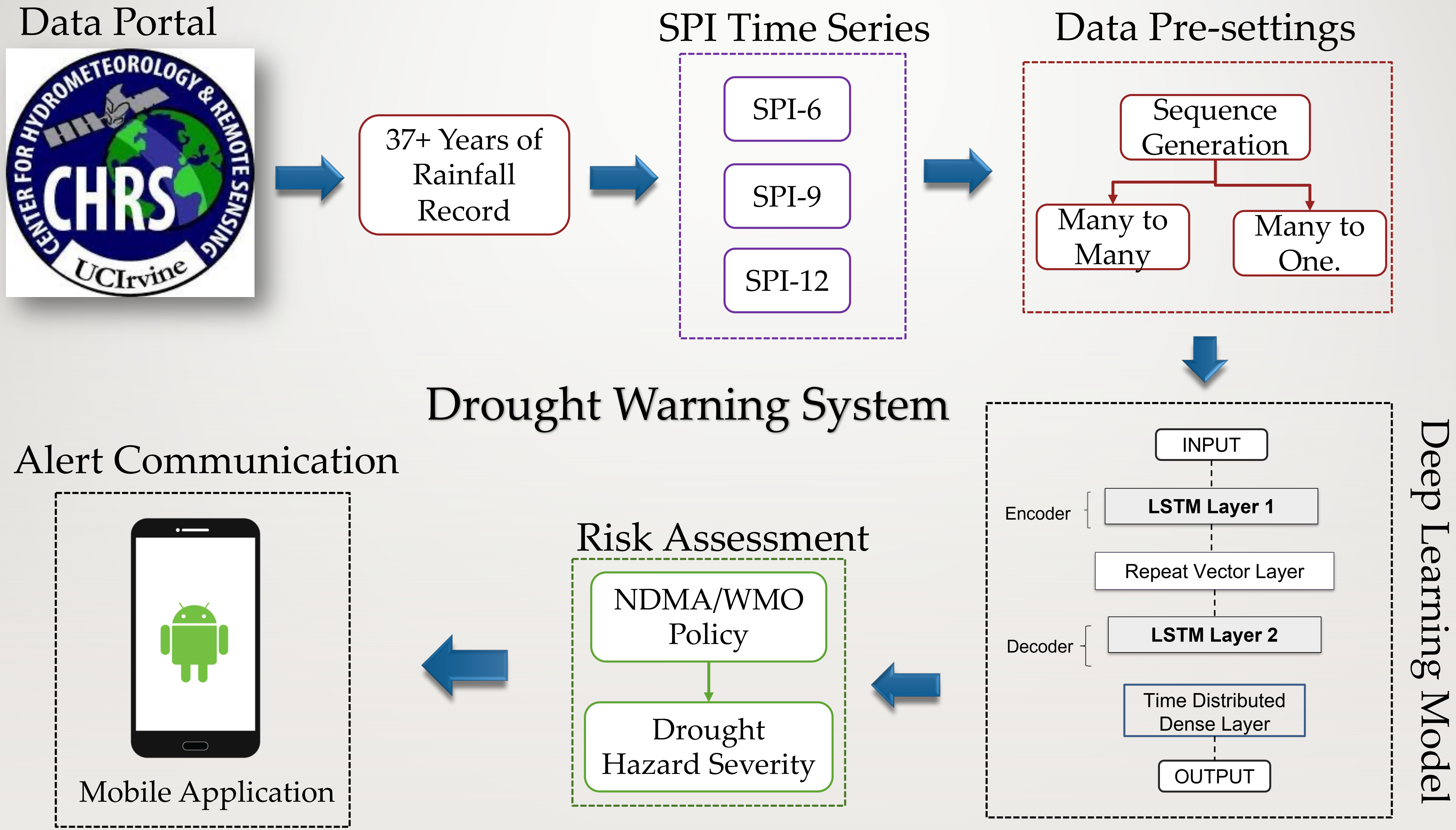
Ongoing Work

- The forecasts from this information can be used to develop the Reservoir Optimization Model.
- The system is to be deployed in real time by interfacing with the sensors database.
- More sophisticated deep learning approaches are to be implemented to make the forecast:
  - More reliable
  - Robust
  - Generalizable

Standardize Precipitation Index (SPI)

- Standardize Precipitation Index is a standard index used to characterize the drought over different timescales.
- SPI over longer time scales are used to model the reservoir storage and ground water.




Data Portal
CENTER FOR HYDROMETEOROLOGY & REMOTE SENSING
CHRS
UCIrvine
37+ Years of Rainfall Record
SPI Time Series
SPI-6
SPI-9
SPI-12
Data Pre-settings
Sequence Generation
Many to Many
Many to One.
Drought Warning System
Deep Learning Model
INPUT
Encoder
LSTM Layer 1
Repeat Vector Layer
Decoder
LSTM Layer 2
Time Distributed Dense Layer
OUTPUT
Risk Assessment
NDMA/WMO Policy
Drought Hazard Severity
Alert Communication
Mobile Application

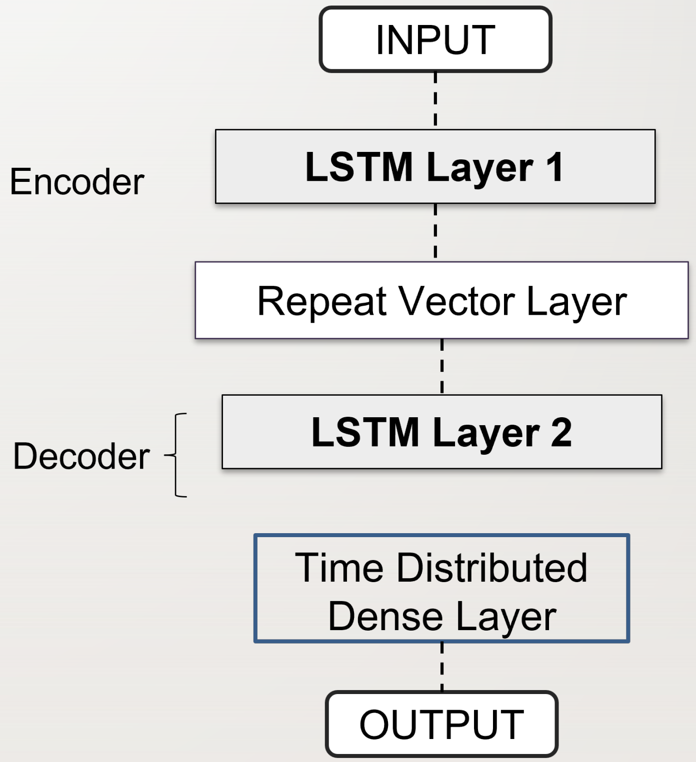
SPI Prediction Model

Training Process	
Number Of LSTM Layers	2
Hidden States In layers	12
Activation	ELU (exponential linear unit)
Batch Size	2
Input	12-Months SPI
Number Of Features	1
Number Of Epochs	100
Optimizer	Adam
Loss Function	Mean Square Error
Output	Next 2 Months Data Points

2-Months Lead Time

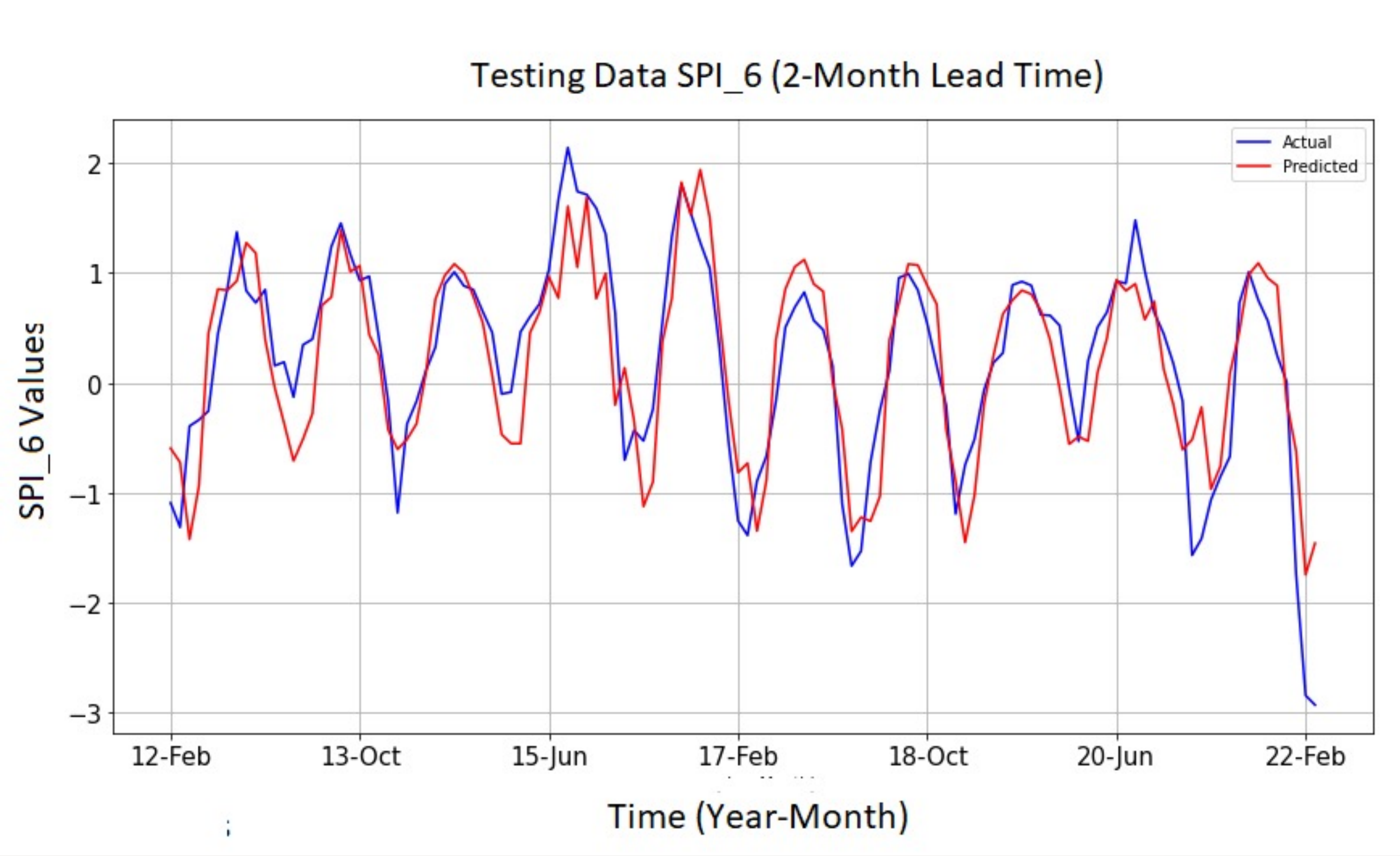
Features

- SPI Time Series



Results

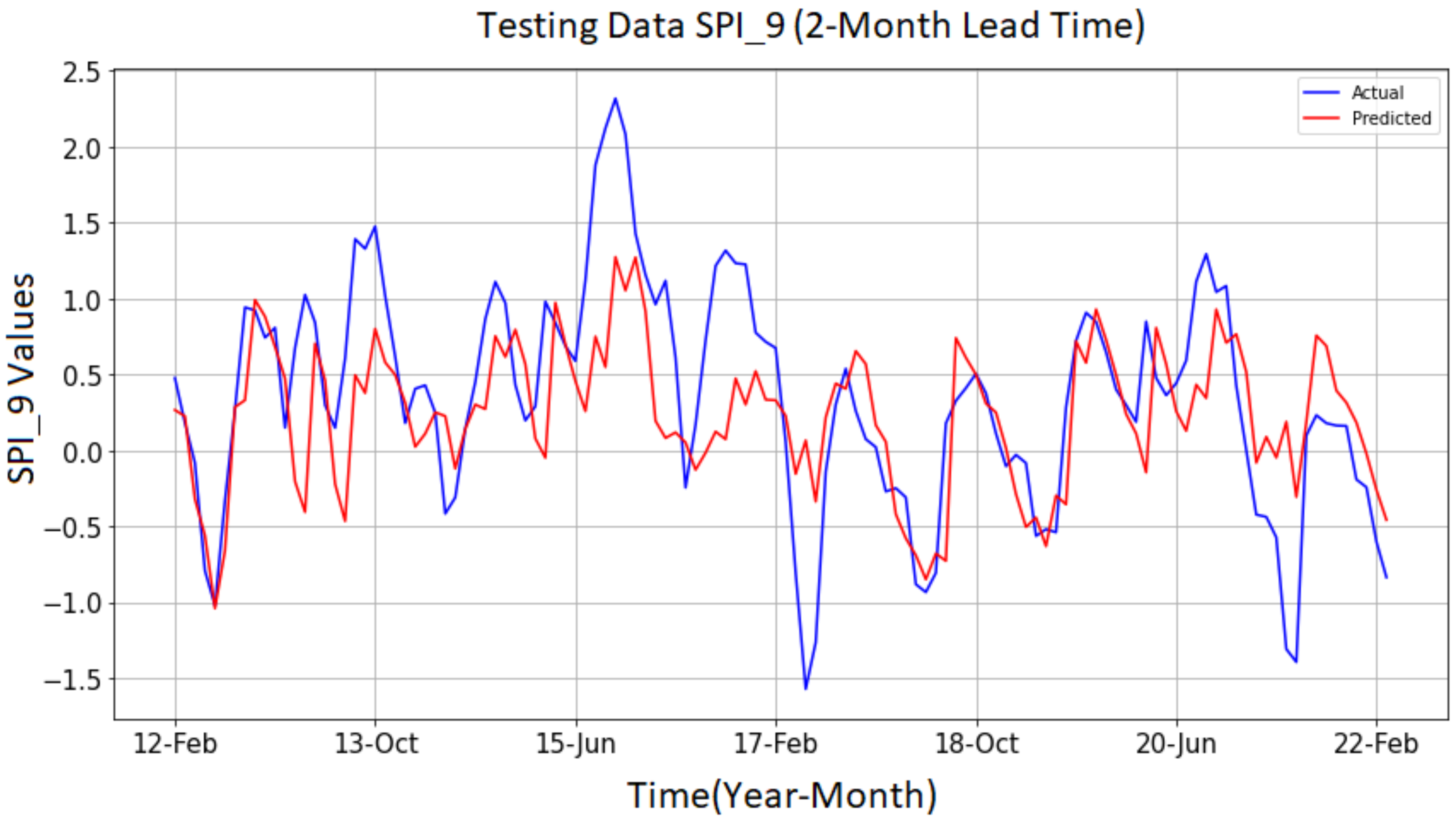
LSTM Input = 1-Year , Prediction = Next 2 Month, **RMSE = 0.4950956**



Year and Month	Predicted SPI-6 Index	Actual SPI-6 Index
Feb 2022	-0.6268	-2.83404
March 2022	-1.9465	-2.92053
April 2022	-1.75636	-2.94954

Results

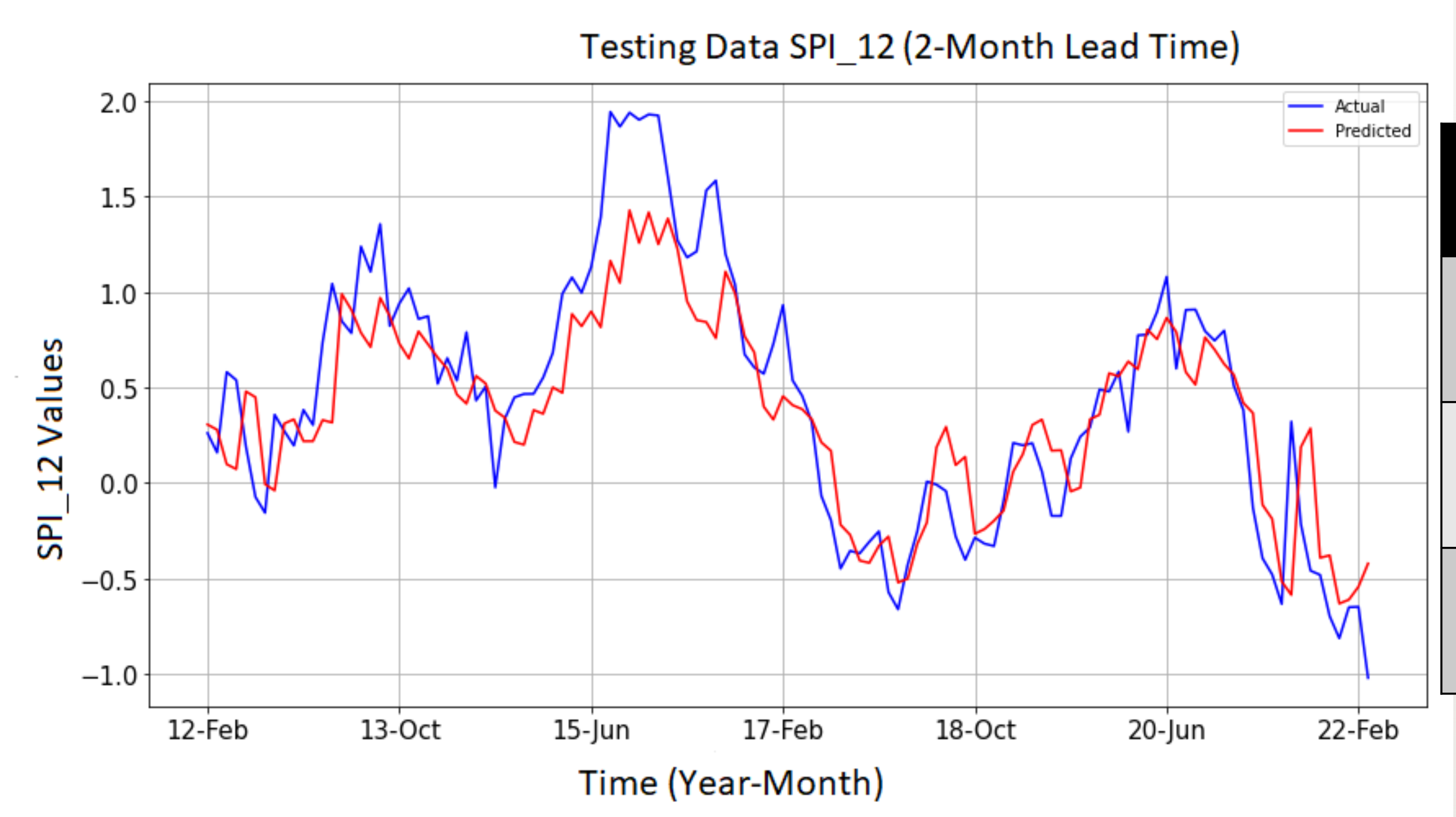
LSTM Input = 1-Year , Prediction = Next 2 Month, **RMSE = 0.566728**



Year and Month	Predicted SPI-9 Index	Actual SPI=9 Index
Feb 2022	-0.2610	-0.58974
March 2022	-0.9345	-0.82241
April 2022	-1.069	-3.11308

Results

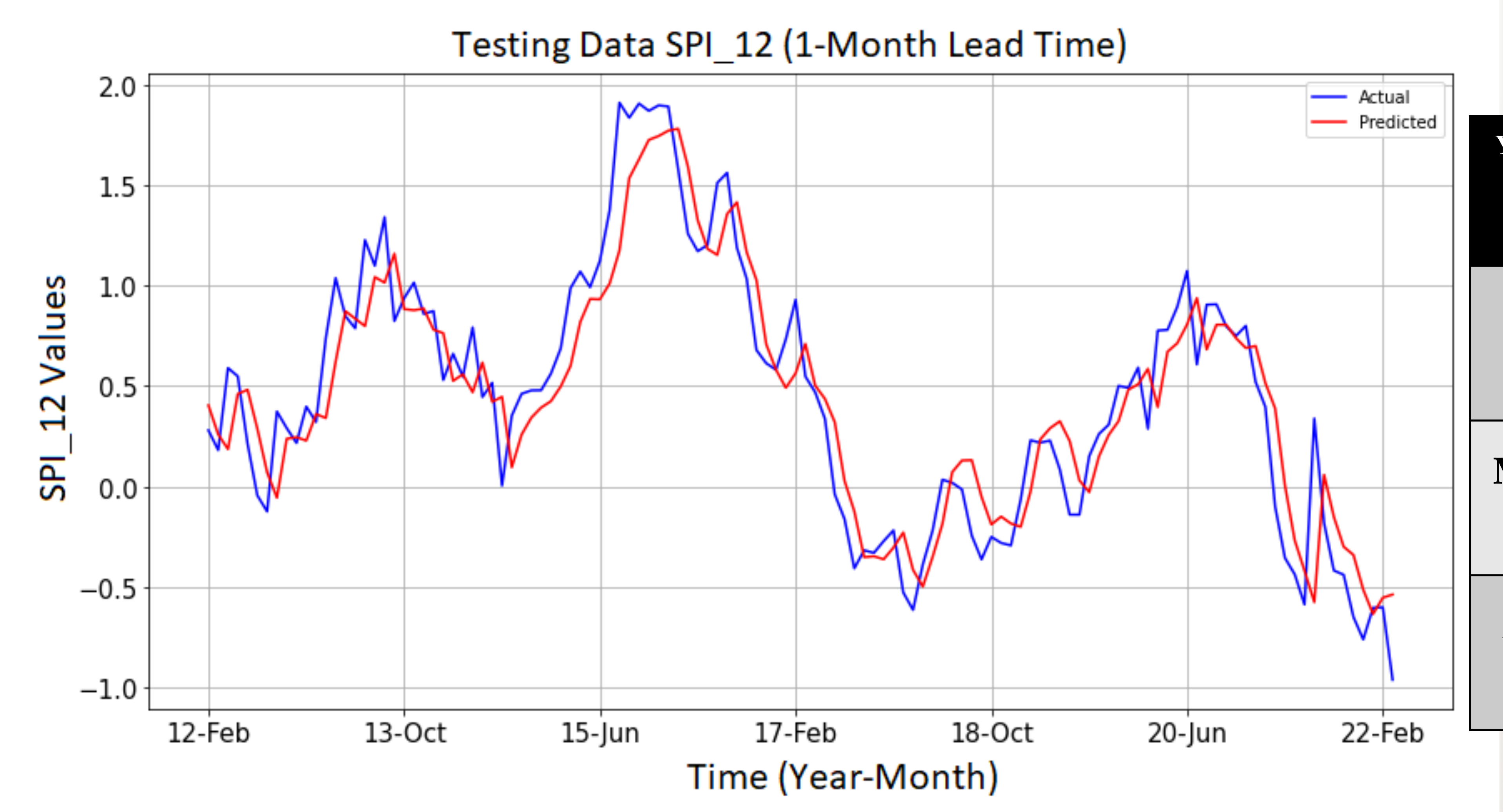
LSTM Input = 1-Year , Prediction = Next 2 Month, **RMSE = 0.325094**



Year and Month	Predicted SPI-12 Index	Actual SPI-12 Index
Feb 2022	-0.543	-0.60149
March 2022	-0.5504	-0.96215
April 2022	-0.9699	-1.21812

Results

LSTM Input = 1-Year , Prediction = Next 1 Month, **RMSE = 0.2444**



Year and Month	Predicted SPI-12 Index	Actual SPI-12 Index
Feb 2022	-0.55352	-0.60149
March 2022	-0.53781	-0.96215
April 2022	-0.86795	-1.21812

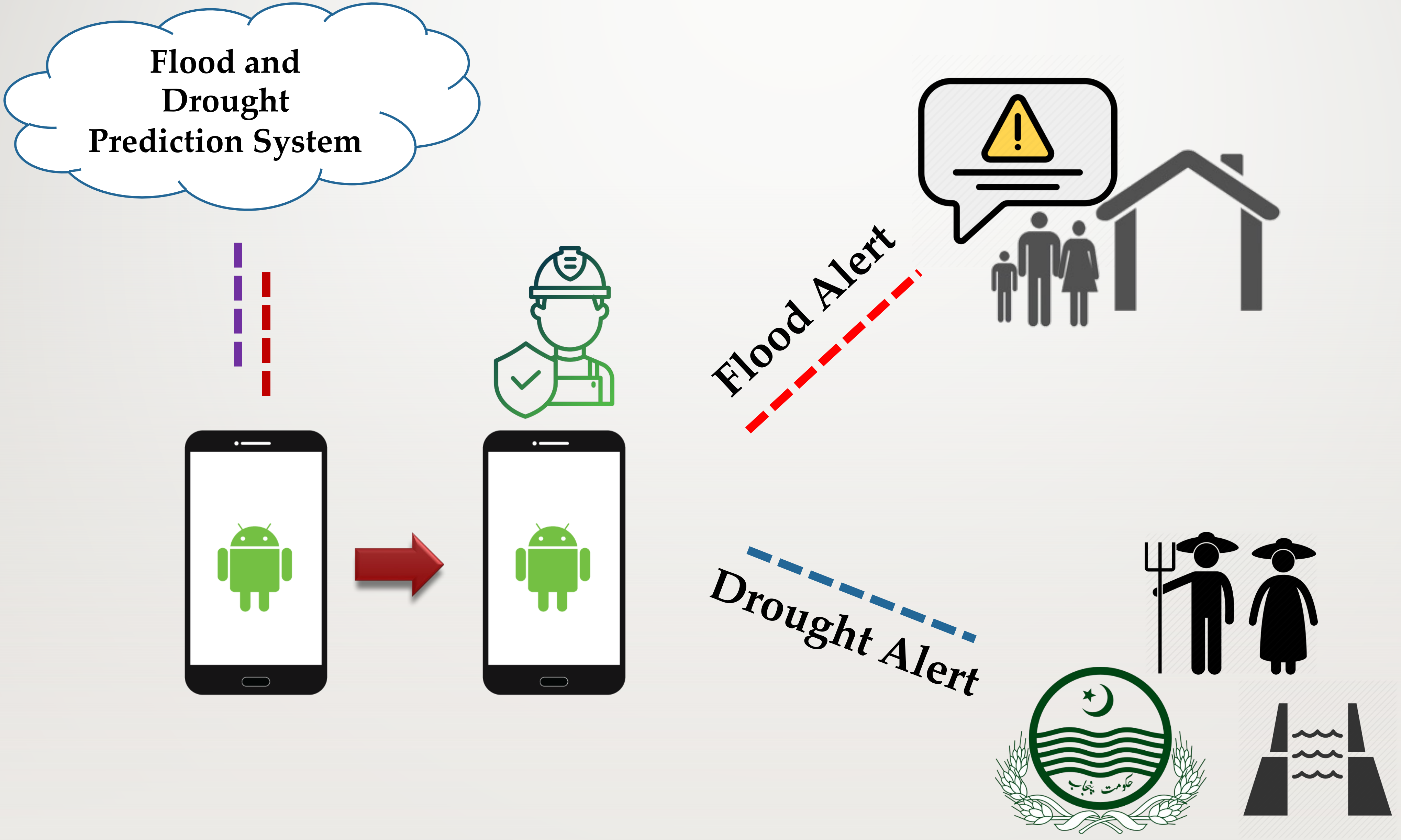
Results Comparison

SPI Prediction Model	RMSE Comparison		
	SPI_6	SPI_9	SPI_12
1-Month Lead Time	0.393155	0.3628945	0.2444
2-Month Lead Time	0.4950956	0.56672787	0.325094

Risk Assessment

Drought Classification for SPI Series	
SPI Value	**Class**
> 0	Normal
0 to -0.55	Mild Drought
-0.5 to -0.84	Moderate Drought
-0.84 to -1.28	Severe Drought
-1.28 to -1.65	Extreme Drought
< -1.65	Very Extreme Drought

Alert Communication



Ongoing Work

- The forecasts from this information can be used to develop the Reservoir Optimization Model.
- The system is to be deployed in real time by interfacing with the sensors database.
- More sophisticated deep learning approaches are to be implemented to make the forecast:
  - More reliable
  - Robust
  - Generalizable

Thank You ☺

jahan@lums.edu.pk