Intro to Classification Networks

Kathy Breen

April 26, 2019

Difference between classification and regression tasks

- **Regression:** approximate the output of a function *(y)* for each input *(x)* in the domain.
  - $y = x^2$
  - $C_{max} = \frac{C_0 A}{4t\pi(D_x D_y)^{1/2}}$
- **Classification:** approximate labels to inputs which sort data into different classes (i.e. categories/types).
  - Land-use type
  - Is this a picture of a cat?

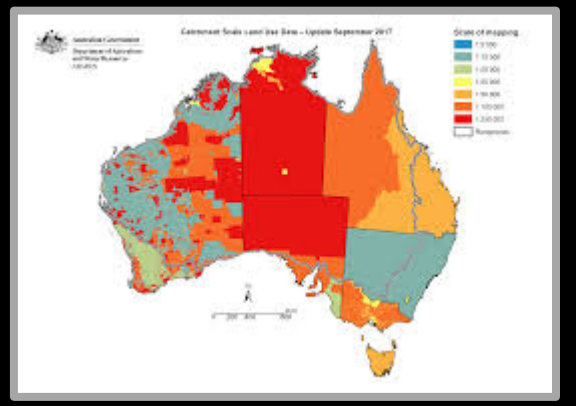
Supervised vs. unsupervised classification

- **Unsupervised:** Statistically based methods that require only inputs to create classes within data. Often exploratory.
  - Principle Components Analysis
  - K-means clustering

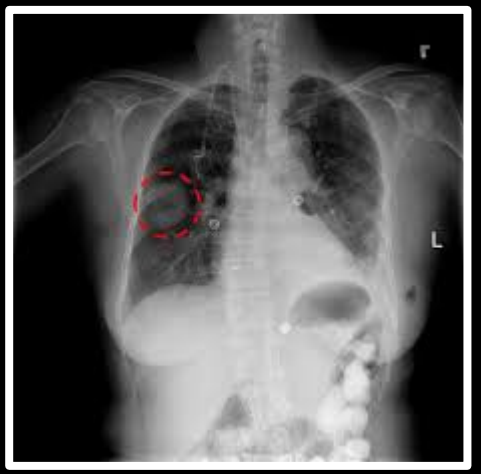
- **Supervised:** Weighted linear combination of inputs are iteratively trained to predict desired (known) target data.
  - Neural networks
  - Boosted regression trees
  - Random forests

Common classification tasks

- **Image Classification:** determine whether an image contains a particular object
  - EX: Is this a cat or not a cat? (Thank you Andrew Ng)



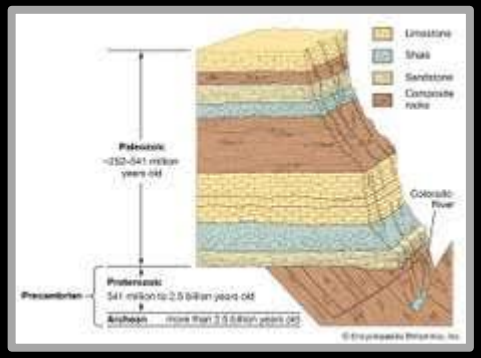
- **Image segmentation:** partition image pixels into multiple classes
  - EX: Land-use classification



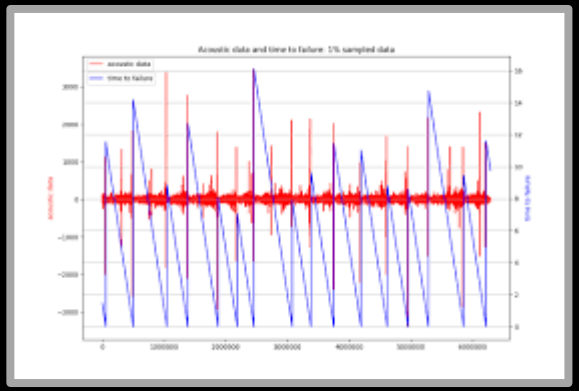
- **Semantic (pixel-based) segmentation:** Determine if a group of pixels represent a target class
  - EX: medical diagnostic imagery

Classification in the Geosciences

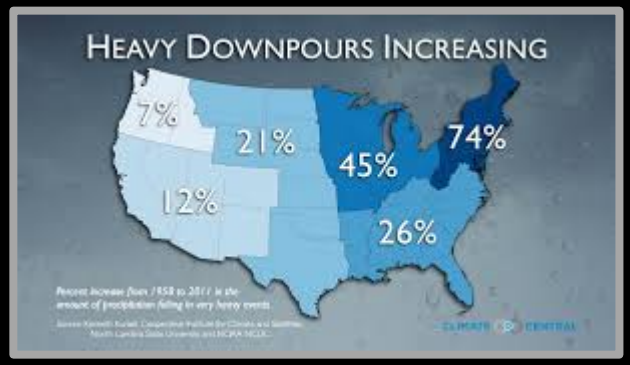
- **Facies Classification:** Given subsurface data (well logs, seismic), predict geologic facies



- **Time to failure:** Given acoustic signals, predict the time to failure for a fault



- **Extreme event prediction:** Given environmental conditions, predict if an extreme event will occur



Code changes: regression vs. classification

	Regression	Classification
Preprocessing	Not necessary for outputs	Outputs must be transformed as binary features
Final activation function	Any, ReLU is common	Softmax
Loss function	Difference between true values and predictions (RMSE)	Probability that true label will be predicted (Cross entropy)

Categorical Targets

Class Label
1
2
3

1	2	3
1	0	0
0	1	0
0	0	1

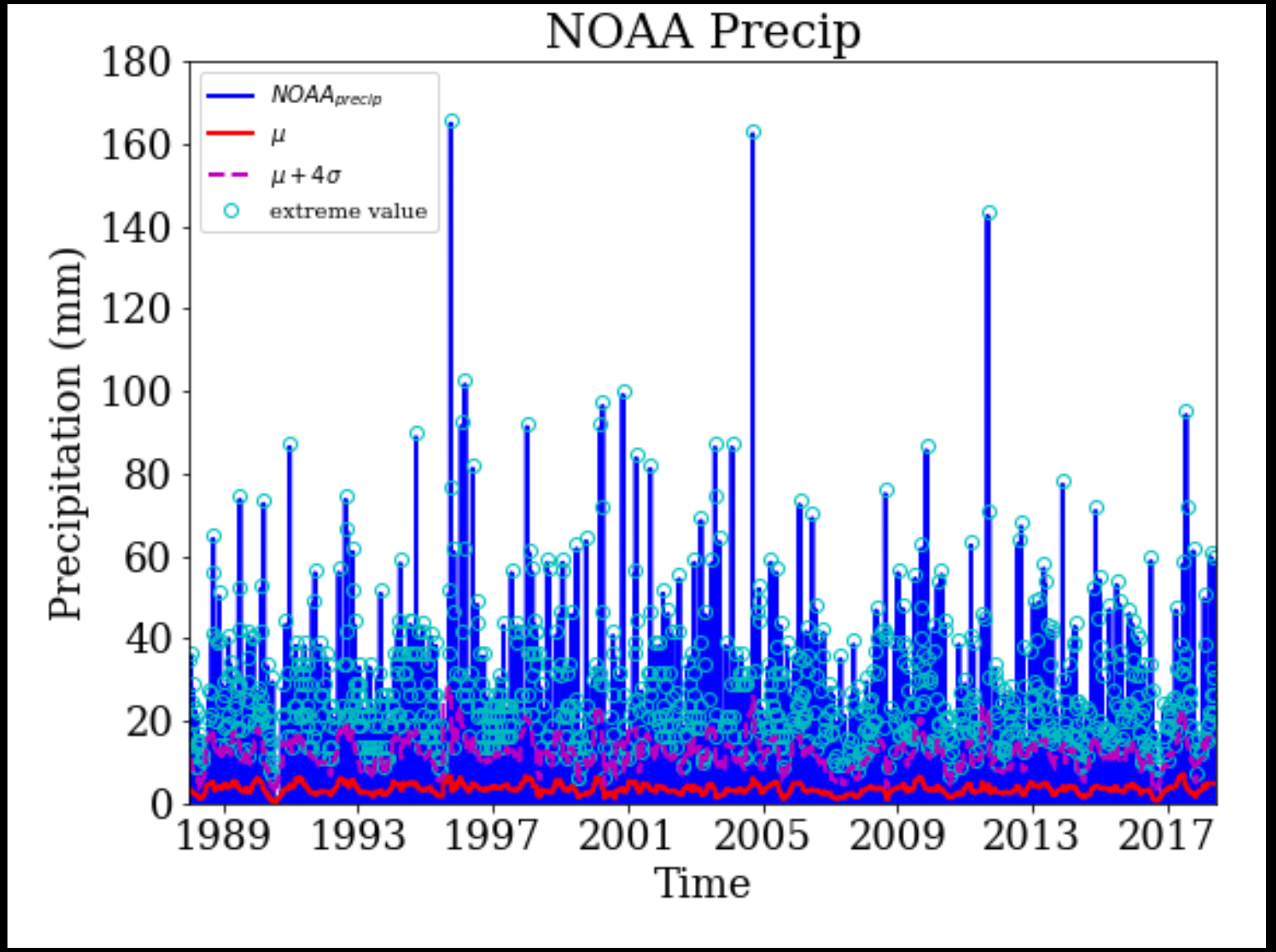
Softmax activation function

The softmax activation function outputs a vector that represents the probability distributions of a list of potential outcomes (i.e. probability that a class was predicted accurately).

Cross Entropy

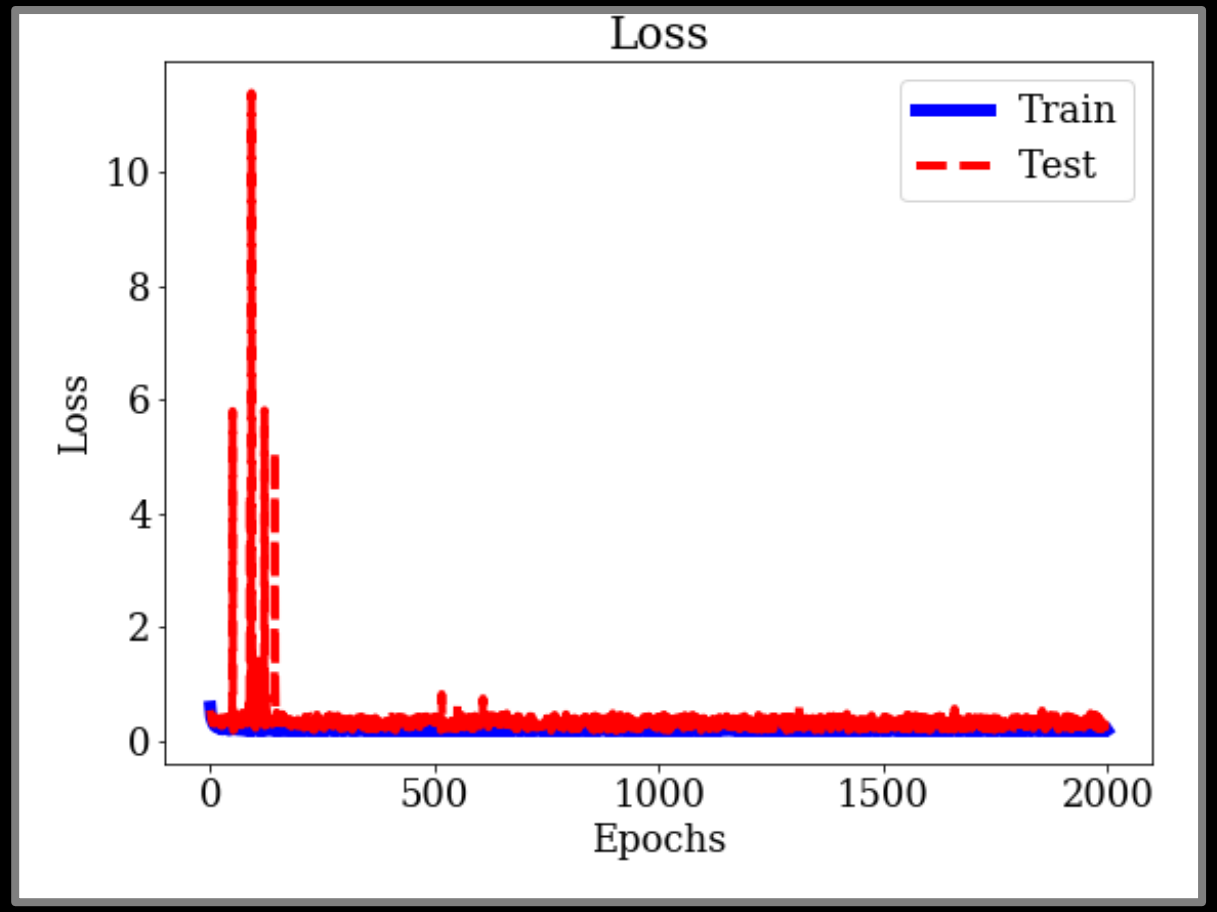
- **Targets have been transformed as binary features (0 or 1)**
- **Softmax outputs probabilities for each target ranging from 0 to 1**
- **Cross entropy loss functions measure how close the predicted probability is to the true class label**
  - Increases as predicted probability (softmax) diverges from the true label (loss function penalized)

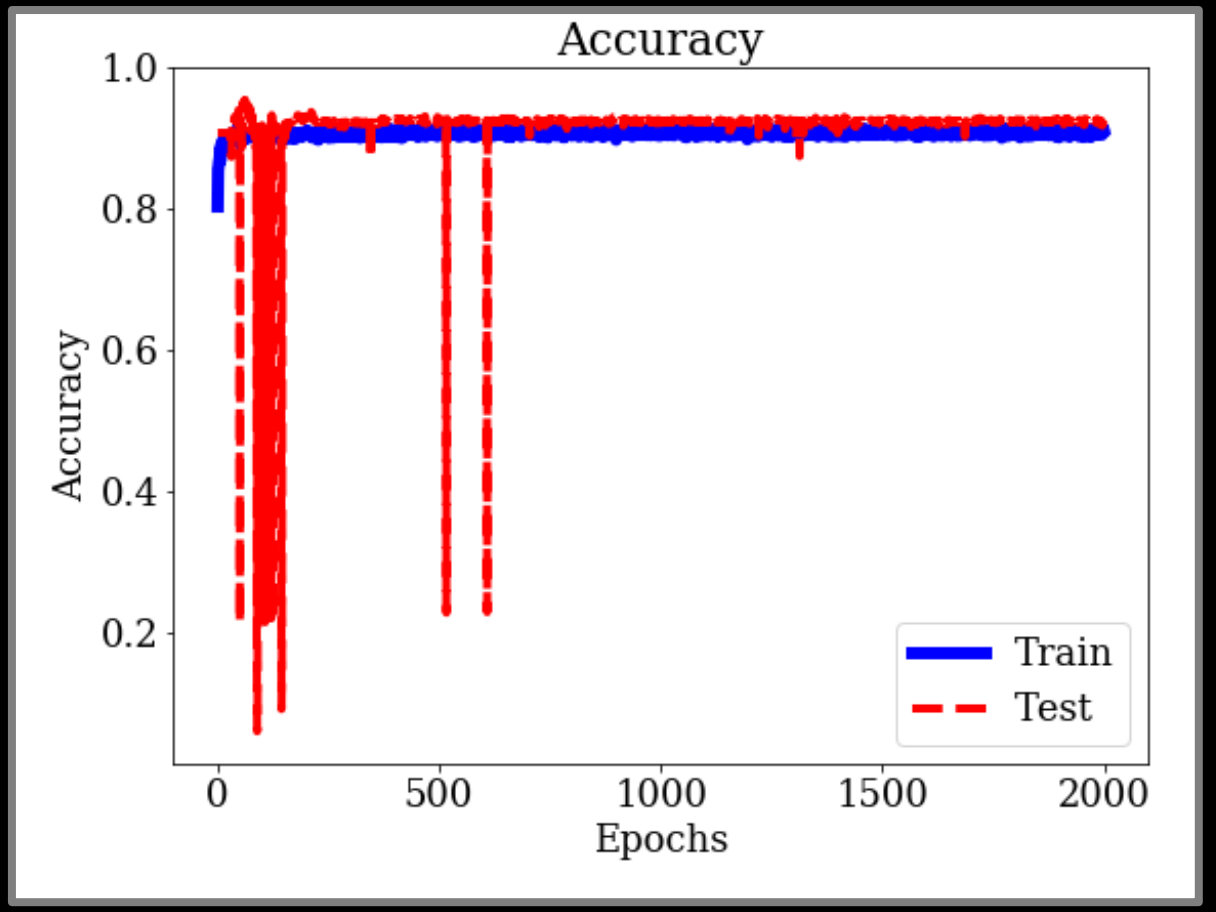
Demo: Binary classification



DEMO

Demo: Results

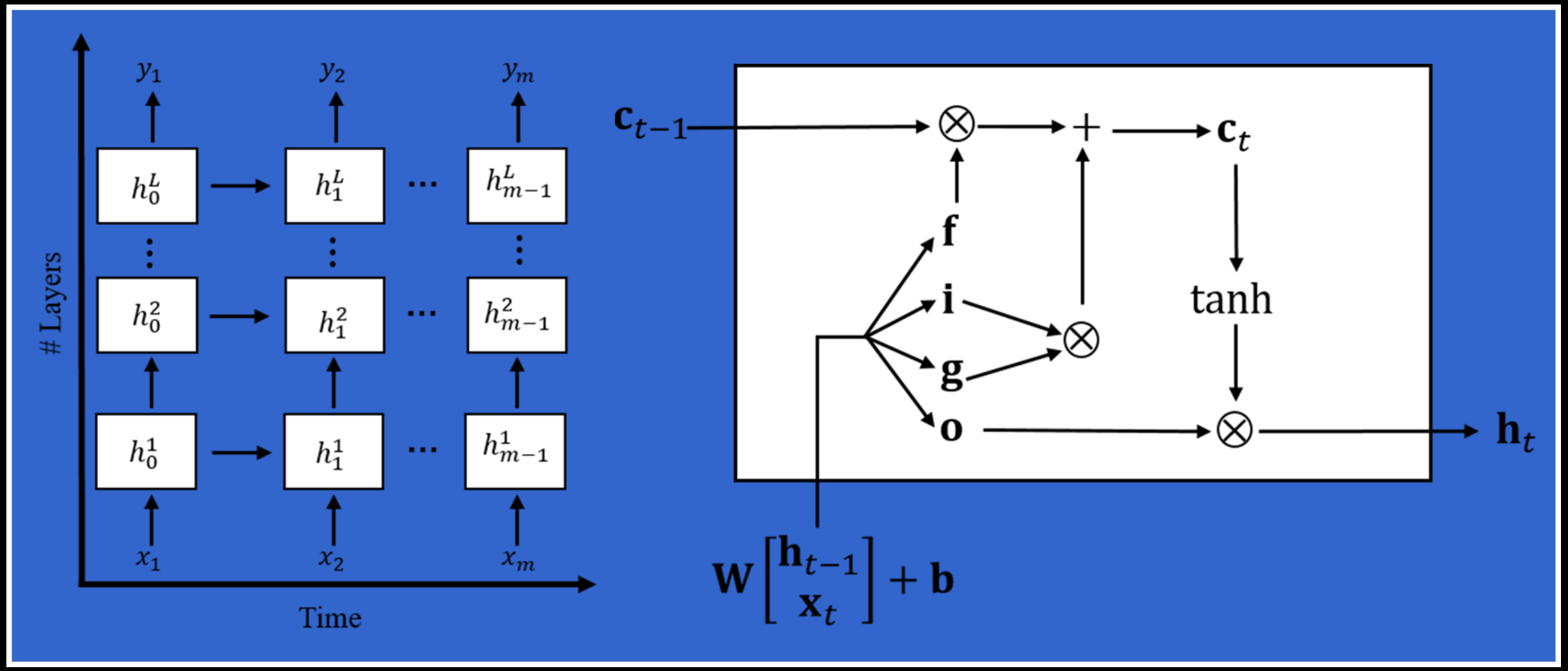




precip_DLM.py: Test Accuracy 93%

Demo: How can results be improved?

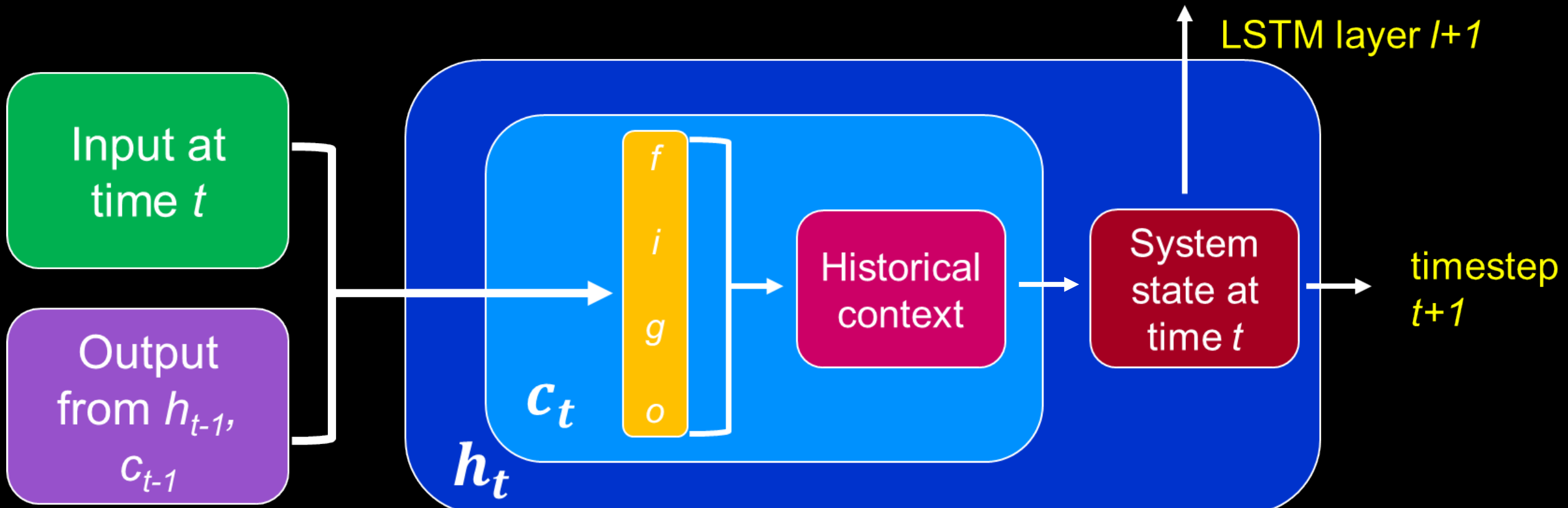
- More data
- Changes in preprocessing (i.e. type of norm, feature engineering) Are we asking a question that is answerable with the data at hand?
- How can we improve the way we formulate the question of interest?
- Architecture appropriate for conceptual problem



Recurrent Neural Network (RNN)

Long Short-Term Memory (LSTM) cell

LSTM Cell


LSTM layer *l+1*
Input at time *t*
Output from h_{t-1}, c_{t-1}
h_t
c_t
f
i
g
o
Historical context
System state at time *t*
timestep *t+1*