Crime Hot Spot Forecasting:
A Recurrent Model with Spatial and Temporal Information

Yong Zhuang[1] Matthew Almeida[1]
Mentors: Wei Ding[1], Melissa Morabito[2]

[1]University of Massachusetts Boston

[2]University of Massachusetts Lowell

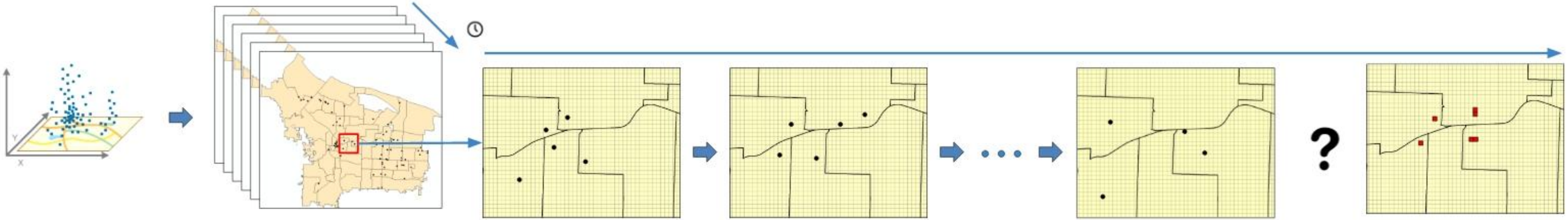
Agenda

- Introduction to crime hotspots forecasting project
- Spatio-Temporal data analysis
- Potential Hot Cells Selection
- Proposed Model
- Experiments and Results
- Conclusion and Future Work
- Acknowledgments

Agenda

- ❖ Introduction to crime hotspots forecasting project
- ❖ Spatio-Temporal data analysis
- ❖ Potential Hot Cells Selection
- ❖ Proposed Model
- ❖ Experiments and Results
- ❖ Conclusion and Future Work
- ❖ Acknowledgments

Crime hotspots forecasting



Understanding patterns in criminal activity allows for the prediction of future crime and enables police precincts to more effectively allocate officers to prevent or respond to incidents.

- **Goal:** Design a predictive model to identify high-risk "hot spots" in the near future based upon the historical neighborhood crime information of the potential hot region.

- **Data:** The call-for-service data provided by the Portland, Oregon Police Bureau (PPB) for a 5-year period from March 2012 through the end of December 2016

Calls-for-service (CFS) records

The CFS data-set includes location information (x_coordinate, y_coordinate), time information (occ_date) and class label (category).

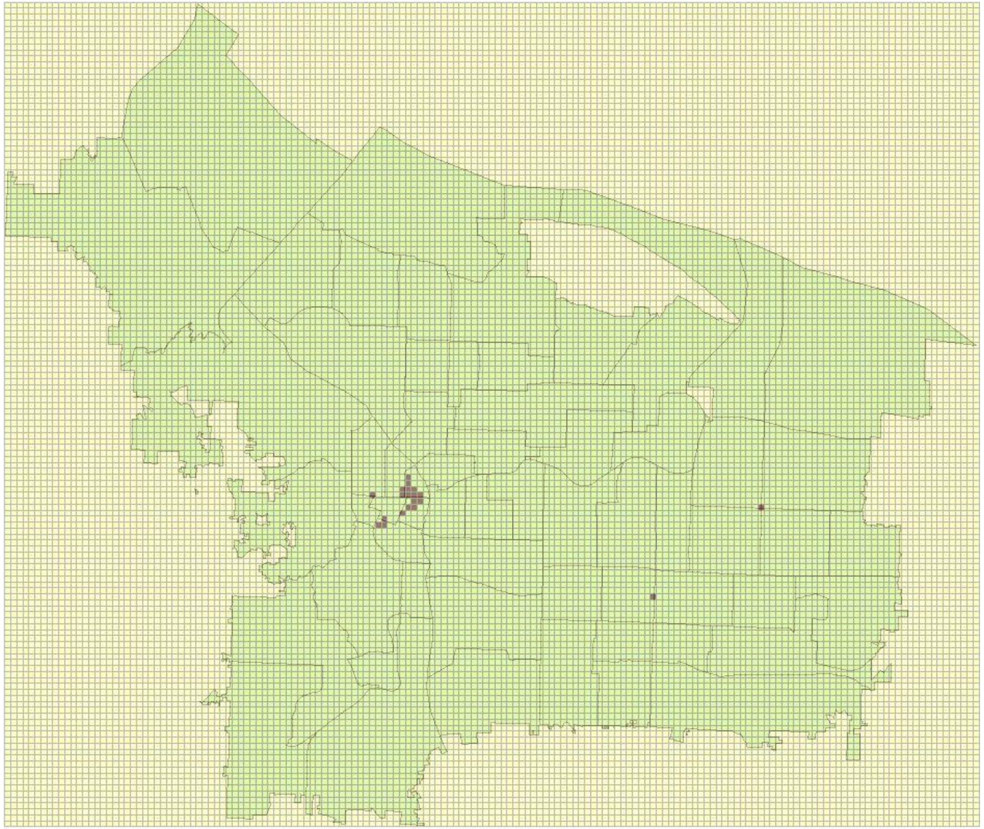
CATEGORY	CALL GROUPS	final_case_type	CASE DESC	occ_date	x_coordinate	y_coordinate	census_tract
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	1/18/2013	7649793	662388	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	1/5/2013	7651202	661479	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	1/28/2013	7647818	663182	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	2/2/2013	7649298	661246	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	1/13/2013	7650935	661746	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	2/17/2013	7650248	660907	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	1/30/2013	7650289	662464	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	3/13/2013	7650182	664208	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	2/16/2013	7649859	665351	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	3/2/2013	7649894	664127	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	3/29/2013	7649298	661246	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	4/27/2013	7647366	665494	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	4/27/2013	7648668	662094	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	5/2/2013	7650785	661371	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	5/12/2013	7647366	665494	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	5/31/2013	7650022	663852	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	6/1/2013	7648386	663997	100
STREET CRIMES	DISORDER	DISTP	DISTURBANCE - PRIORITY	5/27/2013	7648851	662894	100

UMASS BOSTON

Hot Spots Map

Crime hotspot maps are a well-established tool for visualization of space–time crime patterns and can be used as a method for prediction of near-repeat crimes.

Hot Spots Map of Portland

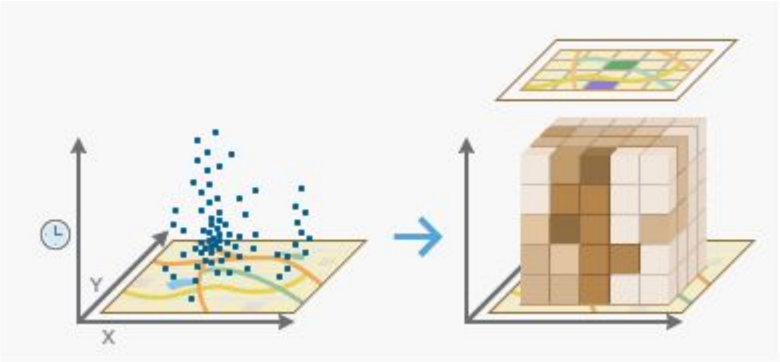


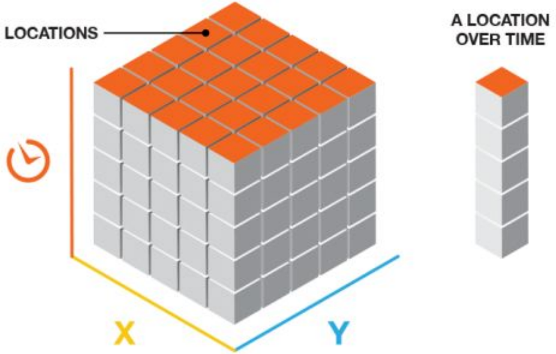
Agenda

- ❖ Introduction to crime hotspots forecasting project
- ❖ Spatio-Temporal data analysis
- ❖ Potential Hot Cells Selection
- ❖ Proposed Model
- ❖ Experiments and Results
- ❖ Conclusion and Future Work
- ❖ Acknowledgments

Spatio-Temporal Data

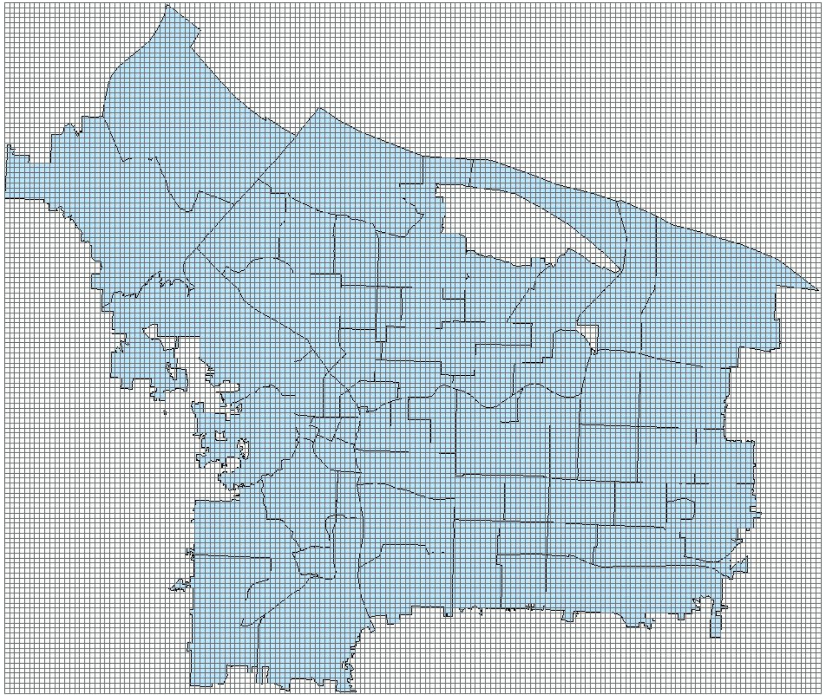
Variables of call-for-services data			
category	occurrence date	x coordinate	y coordinate





Spatial Parameters

Map Size (Portland) : 138 x 163 cells



Cell size 600*600 sq.ft

Spatio-Temporal Analysis

- **Temporal Analysis:** How to Identify the temporal influence?

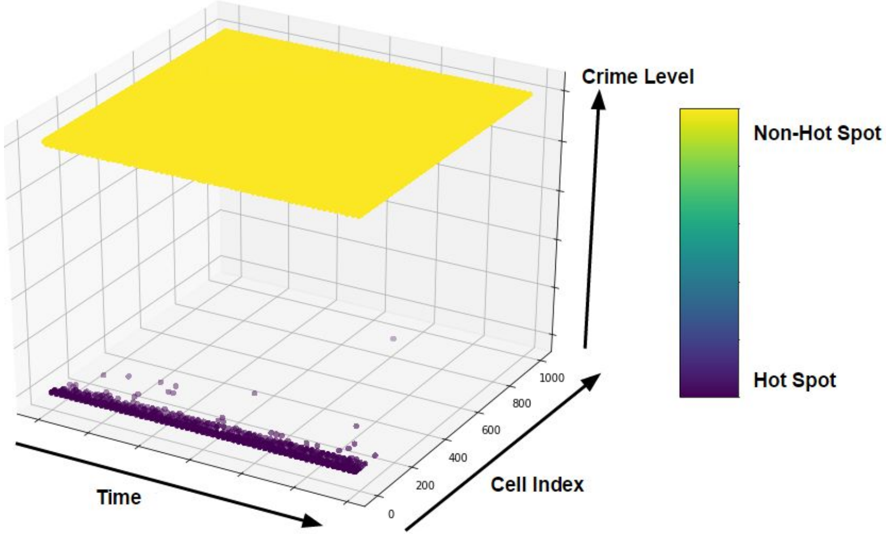
- **Spatial Analysis:** How to use the spatial information?

Agenda

- ❖ Introduction to crime hotspots forecasting project
- ❖ Spatio-Temporal data analysis
- ❖ **Potential Hot Cells Selection**
- ❖ Proposed Model
- ❖ Experiments and Results
- ❖ Conclusion and Future Work
- ❖ Acknowledgments

Potential Hot Cells Selection

Mobility of the hot spots: How likely was it that the areas with the highest crime numbers would be the same month-to-month and year-to-year? If socioeconomic factors remain largely constant over the course of the data, it may be difficult for these patterns to change.



The 3D plot to present the ranking of the hottest cells. The two horizontal axes show the time index and cell index. The vertical axis shows a tiered measure of crime activity and reports whether or not that cell-interval was hot (0) or not(1). It can be seen the ranking was veryresistant to change.

Agenda

- ❖ Introduction to crime hotspots forecasting project
- ❖ Spatio-Temporal data analysis
- ❖ Potential Hot Cells Selection
- ❖ **Proposed Model**
- ❖ Experiments and Results
- ❖ Conclusion and Future Work
- ❖ Acknowledgments

UMASS BOSTON

Proposed Model

❖ **Temporal Analysis:** **Recurrent Neural Network**

Generally, a recurrent neural network is able to combine the input and the hidden state of one step with the inner weight matrices to generate the hidden state on the next step. The hidden state can be computed as follows:

$$h^s_{t_i} = f(a \cdot m^s_{t_i} + b \cdot h^s_{t_{i-1}}) \tag{1}$$

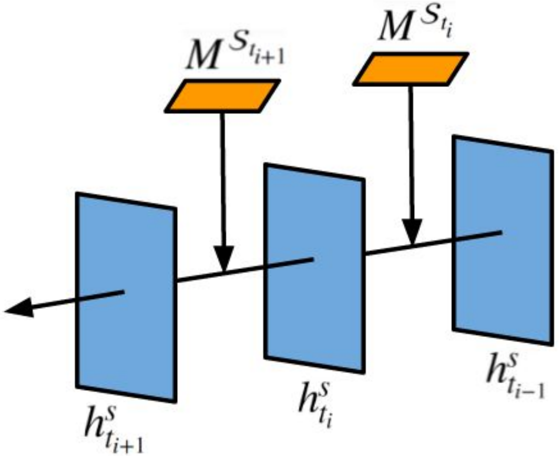
where $h^s_{t_i}$ presents the historical information of a cell s until the time interval t_i, $m^s_{t_i}$ denotes the number of crimes occurred in s at time interval t_i. a and b are the weights. And the activation function $f(x)$ is chosen as a rectifier function as follows:

$$h^s_{t_i} = \max\{a \cdot m^s_{t_i} + b \cdot h^s_{t_{i-1}}, 0\} \tag{2}$$

UMASS BOSTON

Proposed Model

❖ Spatio-Temporal Analysis:



Spatio-Temporal Neural Networks

To build a precise space-time series analysis model, both temporal and spatial influence should be considered. In each cell, both its own crime count and that of the cells around it are essential factors for predicting the crime level of the cell, so it is necessary to include all of that information in our model. Therefore the historical spacial influence can be considered as the accumulation of the short-term influences of recent, nearby events. Since influence of a certain event will decrease with the extension of time and space, the historical spacial influence at a cell s can be limited to only those events that have influence that can "reach" that cell, those that are in a particular space-time window $\mathcal{S}_t \subset S$, where $\mathcal{S}_t = \{[t-\Delta t, t], [x_s - \Delta x, x_s + \Delta x], [y_s - \Delta y, y_s + \Delta y]\}$. As shown in Figure 4, given a cell s, its crime level at a specific time t can be denoted as follows:

$$h^s_{t_i} = f(a \cdot M^{\mathcal{S}_{t_i}} + b \cdot h^s_{t_{i-1}}) \tag{3}$$

Here, $M^{\mathcal{S}_{t_i}}$ computes the total spatial influence at cell s in time interval t_i, and it can be described as:

$$M^{\mathcal{S}_{t_i}} = \bigcup_{x_k, y_k \in \mathcal{S}_{t_i}} m^k_{t_i} \tag{4}$$

Architect of Spatio-Temporal Neural Networks (STNN)

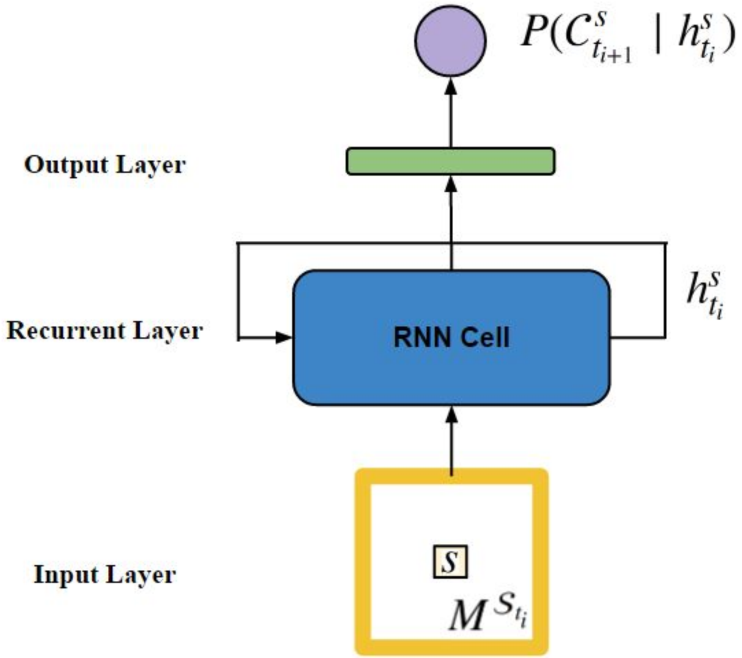
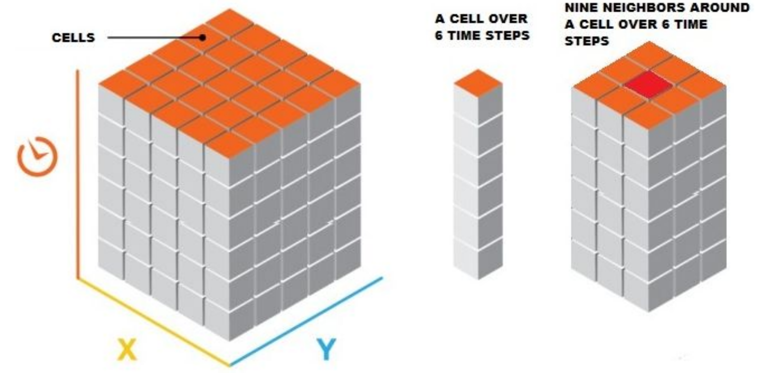


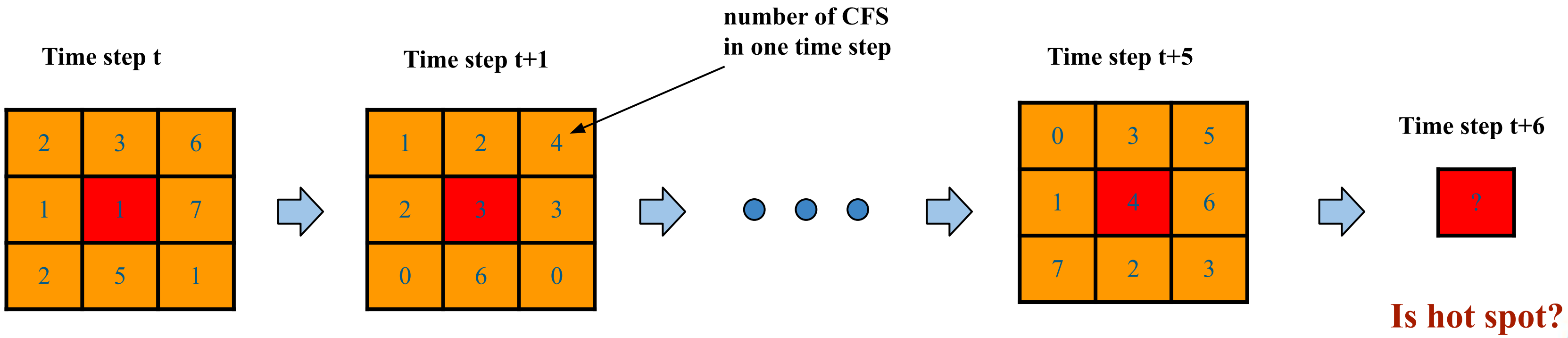
Figure 5. Architect of STNN. For a given sequence M^s, the recurrent layer learns a representation that summaries the nonlinear dependency over the previous events. Based on the learned hidden state h^s_t , it outputs the prediction crime level of cell s on next time step.

UMASS BOSTON

Sample

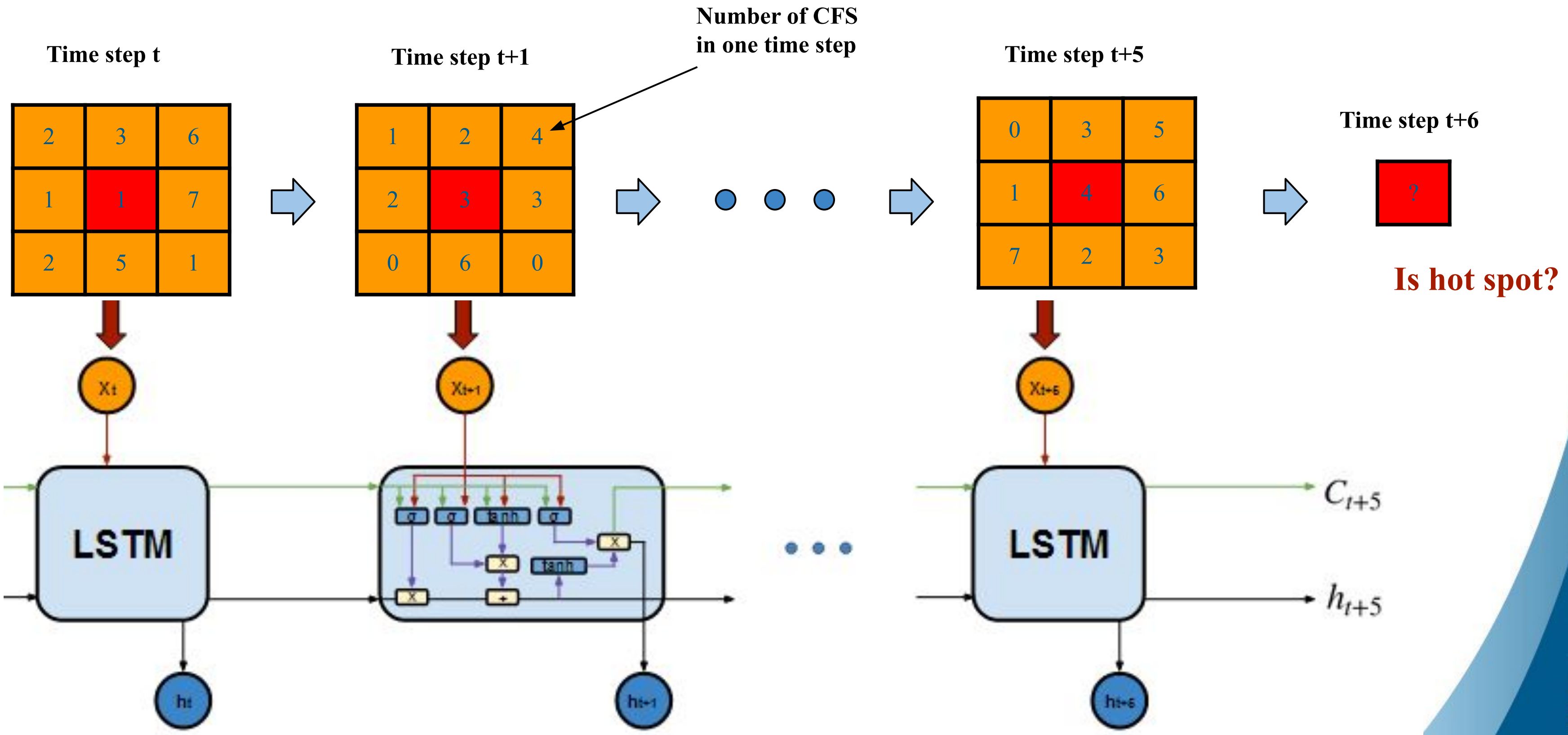


CELLS
X
Y
A CELL OVER
6 TIME STEPS
NINE NEIGHBORS AROUND
A CELL OVER 6 TIME
STEPS

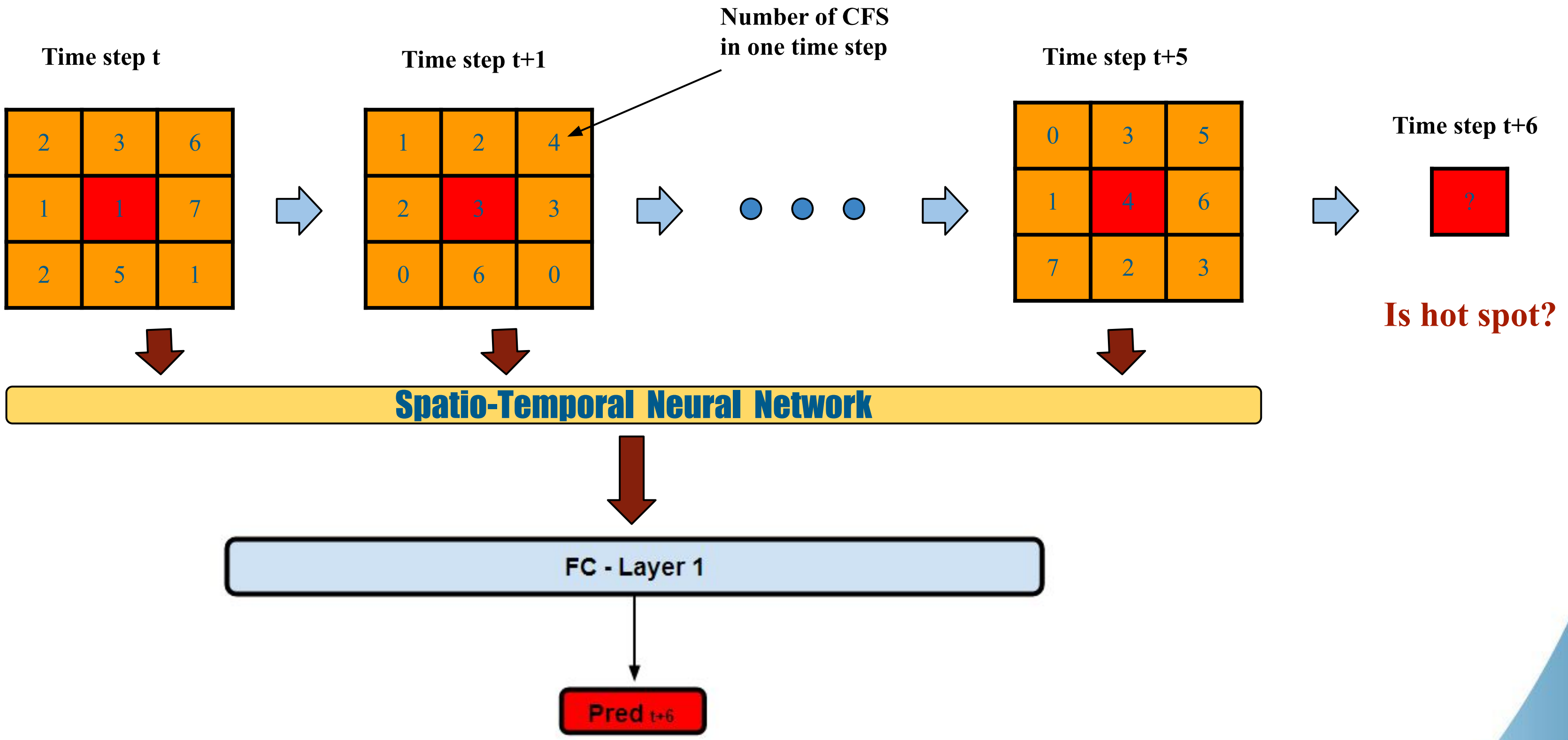


Time step t
2 3 6
1 1 7
2 5 1
Time step t+1
1 2 4
2 3 3
0 6 0
number of CFS
in one time step
Time step t+5
0 3 5
1 4 6
7 2 3
Time step t+6
?
Is hot spot?

Spatio-Temporal Neural Network



Hot Spot Forecasting Network



Agenda

- ❖ Introduction to crime hotspots forecasting project
- ❖ Spatio-Temporal data analysis
- ❖ Potential Hot Cells Selection
- ❖ Proposed Model
- ❖ Experiments and Results
- ❖ Conclusion and Future Work
- ❖ Acknowledgments

UMASS BOSTON

Experiments and Results

- **Comparisons with Three Options for the RNN Architecture**
  - Recurrent Neural Network (RNN)
  - Long Short-Term Memory (LSTM)
  - Gated Recurrent Unit (GRU)

TABLE 2. PERFORMANCE COMPARISON ON 3 RNN ARCHITECTURES EVALUATED BY ACCURACY, PRECISION, RECALL, AND F1-SCORE.

Models	Accuracy	Precision	Recall	F1-score
STNN –RNN	0.75	0.865	0.703	0.776
STNN –LSTM	0.815	0.87	0.745	0.801
STNN –GRU	0.815	0.862	0.75	0.802

TABLE 1. THE PARAMETERS FOR EXPERIMENTS ON STNN.

Subject	Parameter	Value
Network	Learning Rate	0.0005
	Input Neurons	441
	Memory Neurons	20
	Time Steps	24
	Number of Class	2
Training	Positive Samples	1,700
	Negative Samples	1,700
	Training Epochs	300
Validation	Positive Samples	200
	Negative Samples	200
Testing	Positive Samples	200
	Negative Samples	200

Experiments and Results

❖ Comparisons with Traditional Classification Algorithms

TABLE 3. PERFORMANCE COMPARISON ON STNN (LSTM) AND 6 BASELINE METHODS EVALUATED BY ACCURACY, PRECISION, RECALL,AND F1-SCORE.

Models	Accuracy	Precision	Recall	F1-score	Parameters		
STNN –LSTM	0.815	0.87	0.745	0.801	Memory Neurons 20	Learning Rate 0.0005	Activation relu
Decision Tree	0.76	0.806	0.685	0.74	Criterion gini	Max Depth 5	
Gaussian Naive Bayes	0.743	0.79	0.66	0.719			
Random Forests	0.7625	0.803	0.695	0.745	Estimators 10	Min Samples Split 2	
K-nearest neighbors	0.6375	0.71	0.62	0.662	K 1	Distance Measure L2	
Logistic Regression	0.75	0.767	0.725	0.746	Training Epochs 300	penalty L2	
Multi-layer Perceptron	0.7675	0.77	0.766	0.768	Hidden Layer Size (100,50)	Learning Rate 0.001	Activation relu

Agenda

- ❖ Introduction to crime hotspots forecasting project
- ❖ Spatio-Temporal data analysis
- ❖ Potential Hot Cells Selection
- ❖ Proposed Model
- ❖ Experiments and Results
- ❖ Conclusion and Future Work
- ❖ Acknowledgments

UMASS BOSTON

Conclusion

In this paper, we formulate the problem of crime forecasting as a space-time series prediction problem and implement a corresponding deep recurrent neural network with spatial influence embedding to estimate the crime level in the near future and show that it outperforms conventional machine learning algorithms and alternate choices of architecture.

Agenda

- ❖ Introduction to crime hotspots forecasting project
- ❖ Spatio-Temporal data analysis
- ❖ Potential Hot Cells Selection
- ❖ Proposed Model
- ❖ Experiments and Results
- ❖ Conclusion and Future Work
- ❖ Acknowledgments

Acknowledgments

Funding for this research was provided in part by the Oracle Education Foundation grant to the College of Science and Mathematics at the University of Massachusetts Boston.

Thank You!