Neural Message Passing for Multi-Label Classification

Jack Lanchantin, Arshdeep Sekhon, Yanjun Qi

University of Virginia, Department of Computer Science


Castle


Mountains
Castle
Road

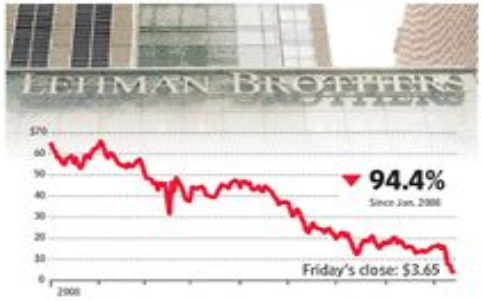
THE WALL STREET JOURNAL.

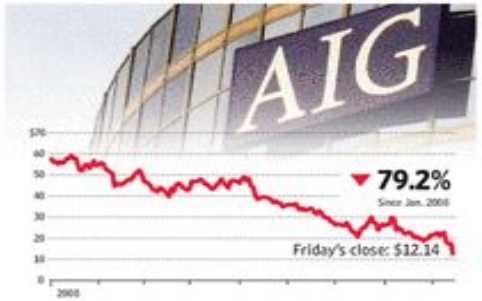
DOWJONES
A News Corporation Company

MONDAY, SEPTEMBER 15, 2008 - VOL. CCLII NO. 64

★★ $2.00

Last week: DJIA 11421.99 ▲201.03 1.8% NASDAQ 2261.27 ▲0.2% NIKKEI 12214.76 unch. DJ STOXX 50 2858.68 ▲3.8% 10-YR TREASURY ▼20/32, yield 3.730% OIL $101.18 ▼$5.05 EURO $1.4217 YEN 107.87





Crisis on Wall Street as Lehman Totters, Merrill Seeks Buyer, AIG Hunts for Cash

U.S. Opts to Avoid Lehman Rescue, Stirring a Momentous Weekend for American Finance; Traders Brace for a Chaotic Monday

By Carrick Mollenkamp, Susanne Craig And Serena Ng

The American financial system was shaken to its core on Sunday. Lehman Brothers Holdings Inc. faced the prospect of liquidation and Merrill Lynch & Co. was close to a deal to sell itself to Bank of America Corp.

The U.S. government, which bailed out Fannie Mae and Freddie Mac a week ago and orchestrated the sale of Bear Stearns Cos. to J.P. Morgan & Co. earlier this year, drew a line in the sand with Lehman. It refused to provide a financial backstop to potential buyers.

Without such support, Barclays PLC and Bank of America, the two most interested buyers, walked away. On Sunday night, Bank of America was close to striking a deal to buy Merrill Lynch for about $44 billion, or $29 a share. Lehman was working on a possible bankruptcy filing.

As worries spread across Wall Street that Lehman wouldn't survive, brokerage firms, hedge funds and other traders moved to disentangle themselves from trades with Lehman. When hopes of a potential sale dimmed, a quiet Sunday on Wall Street turned into a mad rush. Executives and traders hurried to their offices or worked their phones to unwind outstanding contracts with Lehman and to gauge their overall exposure.

A sense of foreboding gripped the Street as top executives feared collateral damage from a Lehman liquidation. Attention turned to Merrill Lynch, which boasts the largest force of retail brokers, and to American International Group Inc., the insurance giant. Both firms have seen their stocks get hammered, and their managements spent the weekend trying to come up with plans to reassure the markets.

"Monday will be a day of reckoning for the financial markets," said Carlos Mendez, senior managing director of ICP Capital, a boutique investment firm in New York. On Sunday, he said, "it was like a fire alarm went off and people ran in all directions."

AIG executives spent the weekend trying to raise cash, either from asset sales or a capital infusion from private-equity firms, or both. AIG executives were meeting with regulators to see if they could transfer capital from some of its subsidiaries to the holding company.

Merrill, whose retail brokerage force is the largest in the country and is known as the "thundering herd," quietly engaged in discussions with Bank of America, whose retail bank branches stretch coast to coast. Wall Street executives said the Federal Reserve may be involved in orchestrating the sale, figuring that it was "better to save the relatively healthy patient instead of the dying one," said a lawyer involved in the discussions.

"We are in uncharted waters here," said a top executive of a big bank. "If Merrill can pull off a deal this weekend, that would certainly help."

The U.S. dollar, which had strengthened in the past few weeks, fell against all four of its major rivals on Sunday—the euro, the Swiss franc, the U.K. pound and the Japanese yen.

Some executives involved in the Lehman discussions held out hope that an 11th-hour reprieve would materialize. Under one scenario aimed at limiting the ripple effects of Lehman's demise, a group of about 15 banks were in discussions Sunday to pool about $100 billion, which would be used to buy assets of the battered securities firm, according to one person familiar with the situation. Details were being finalized Sunday night. One possibility is that the Federal Reserve will support the move by opening its wholesale-borrowing window and relaxing collateral requirements for borrowers.

Lehman, a 158-year-old firm *Please turn to page A18*

Ultimatum By Paulson Sparked Frenzied End

After more than a century and a half as an investment house, Lehman Brothers Holdings Inc. tottered on the brink Sunday night. But its latest troubles were set in motion days earlier.

In midweek, Treasury Secre-

By Deborah Solomon, Dennis K. Berman, Susanne Craig and Carrick Mollenkamp

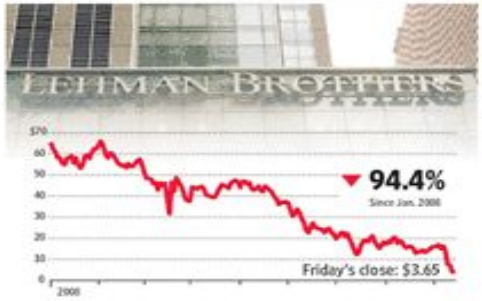
THE WALL STREET JOURNAL.

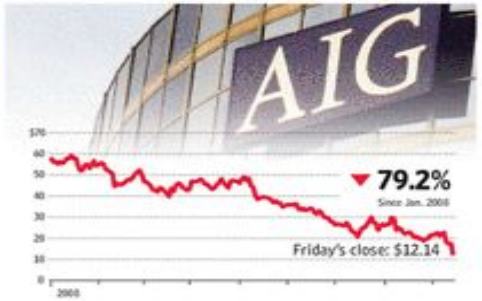
DOWJONES A News Corporation Company

MONDAY, SEPTEMBER 15, 2008 - VOL. CCLII NO. 64

★★ $2.00

Last week: DJIA 11421.99 ▲ 201.03 1.8% NASDAQ 2261.27 ▲ 0.2% NIKKEI 12214.76 unch. DJ STOXX 50 2858.68 ▲ 3.8% 10-YR TREASURY ▼ 20/32, yield 3.730% OIL $101.18 ▼ $5.05 EURO $1.4217 YEN 107.87





Crisis on Wall Street ... an Totters, Merrill Seeks Buyer, ... ts for Cash

Finance

U.S. Opts to Avoid Lehman Rescue, Stirring a Momentous Weekend for American Finance; Traders Brace for a Chaotic Monday

By Carrick Mollenkamp, Susanne Craig And Serena Ng

The American financial system was shaken to its core on Sunday. Lehman Brothers Holdings Inc. faced the prospect of liquidation and Merrill Lynch & Co. was close to a deal to sell itself to Bank of America Corp.

The U.S. government, which bailed out Fannie Mae and Freddie Mac a week ago and orchestrated the sale of Bear Stearns Cos. to J.P. Morgan & Co. earlier this year, drew a line in the sand with Lehman. It refused to provide a financial backstop to potential buyers.

Without such support, Barclays PLC and Bank of America, the two most interested buyers, walked away. On Sunday night, Bank of America was close to striking a deal to buy Merrill Lynch for about $44 billion, or $29 a share. Lehman was working on a possible bankruptcy filing.

As worries spread across Wall Street that Lehman wouldn't survive, brokerage firms, hedge funds and other traders moved to disentangle themselves from trades with Lehman. When hopes of a potential sale dimmed, a quiet Sunday on Wall Street turned into a mad rush. Executives and traders hurried to their offices or worked their phones to unwind outstanding contracts with Lehman and to gauge their overall exposure.

A sense of foreboding gripped the Street as top executives feared collateral damage from a Lehman liquidation. Attention turned to Merrill Lynch, which boasts the largest force of retail brokers, and to American International Group Inc., the insurance giant. Both firms have seen their stocks get hammered, and their managements spent the weekend trying to come up with plans to reassure the markets.

"Monday will be a day of reckoning for the financial markets," said Carlos Mendez, senior managing director of ICP Capital, a boutique investment firm in New York. On Sunday, he said, "it was like a fire alarm went off and people ran in all directions."

AIG executives spent the weekend trying to raise cash, either from asset sales or a capital infusion from private-equity firms, or both. AIG executives were meeting with regulators to see if they could transfer capital from some of its subsidiaries to the holding company.

Merrill, whose retail brokerage force is the largest in the country and is known as the "thundering herd," quietly engaged in discussions with Bank of America, whose retail bank branches stretch coast to coast. Wall Street executives said the Federal Reserve may be involved in orchestrating the sale, figuring that it was "better to save the relatively healthy patient instead of the dying one," said a lawyer involved in the discussions.

"We are in uncharted waters here," said a top executive of a big bank. "If Merrill can pull off a deal this weekend, that would certainly help."

The U.S. dollar, which had strengthened in the past few weeks, fell against all four of its major rivals on Sunday—the euro, the Swiss franc, the U.K. pound and the Japanese yen.

Some executives involved in the Lehman discussions held out hope that an 11th-hour reprieve would materialize. Under one scenario aimed at limiting the ripple effects of Lehman's demise, a group of about 15 banks were in discussions Sunday to pool about $100 billion, which would be used to buy assets of the battered securities firm, according to one person familiar with the situation. Details were being finalized Sunday night. One possibility is that the Federal Reserve will support the move by opening its wholesale-borrowing window and relaxing collateral requirements for borrowers.

Lehman, a 158-year-old firm

Please turn to page A18

Ultimatum By Paulson Sparked Frenzied End

After more than a century and a half as an investment house, Lehman Brothers Holdings Inc. tottered on the brink Sunday night. But its latest troubles were set in motion days earlier.

In midweek, Treasury Secre-

By Deborah Solomon, Dennis K. Berman, Susanne Craig and Carrick Mollenkamp

THE WALL STREET JOURNAL.

DOWJONES
A News Corporation Company

MONDAY, SEPTEMBER 15, 2008 - VOL. CCLII NO. 64

★★ $2.00

Last week: DJIA 11421.99 ▲ 201.03 1.8% NASDAQ 2261.27 ▲ 0.2% NIKKEI 12214.76 unch. DJ STOXX 50 2858.68 ▲ 3.8% 10-YR TREASURY ▼ 20/32, yield 3.730% OIL $101.18 ▼ $5.05 EURO $1.4217 YEN 107.87

LEHMAN BROTHERS

94.4%
Since Jan. 2000
Friday's close: $3.65

Merrill Lynch

68.2%
Since Jan. 2000
Friday's close: $17.05

AIG

79.2%
Since Jan. 2000
Friday's close: $12.14

Crisis on Wall Street ... an Totters, Merrill Seeks Buyer, ... ts for Cash

U.S. Opts ... ing a Momentous Weekend for American Finance; Traders Brace for a Chaotic Monday

By Carrick Mo... Susanne Craig And Serena No...

The American financial system was shaken to its core on Sunday. Lehman Brothers Holdings Inc. faced the prospect of liquidation and Merrill Lynch & Co. was close to a deal to sell itself to Bank of America Corp.

The U.S. government, which bailed out Fannie Mae and Freddie Mac a week ago and orchestrated the sale of Bear Stearns Cos. to J.P. Morgan & Co. earlier this year, drew a line in the sand with Lehman. It refused to provide a financial backstop to potential buyers.

Without such support, Barclays PLC and Bank of America, the two most interested buyers, walked away. On Sunday ... was working on a possible bankruptcy filing.

As worries spread across Wall Street that Lehman wouldn't survive, brokerage firms, hedge funds and other traders moved to disentangle themselves from trades with Leh... man. When hopes ... sale dimmed, a qui... Wall Street turned ... rush. Executives an... ried to their offic... their phones to ums... ling contracts with ... to gauge their over...

A sense of forebo... the Street as top executives feared collateral damage from a Lehman liquidation. Attention ... turned to Merrill Lynch, which boasts the largest force of retail brokers, and to American International Group Inc., the insurance giant. Both firms have seen their stocks get hammered, and their managements spent the weekend trying to come up with plans to reassure the markets.

"Monday will be a day of reckoning for the financial markets," said Carlos Mendez, senior man... ... a ... in ... it ... d ... e ... al ... ity ... firms, or both. AIG executives were meeting with regulators to see if they could transfer capital from some of its subsidiaries to the holding company.

Merrill, whose retail brokerage force is the ... country and ... "thundering ... gaged in disc... of America, ... branches str... Wall Street ... Federal Reser... in orchestrat... ing that it was "better to save the relatively healthy patient instead of the dying one," said a lawyer involved in the discussions.

"We are in uncharted waters here," said a top executive of a big bank. "If Merrill can pull off a deal this weekend, that would certainly help."

The U.S. dollar, which had strengthened in the past few weeks, fell against all four of its major rivals on Sunday—the euro, the Swiss franc, the U.K. pound and the Japanese yen.

... executives involved in ... held out ... reprieve ... one sce... ... he ripple ... mise, a ... were in ... o pool ... h would be used to buy assets of the battered securities firm, according to one person familiar with the situation. Details were being finalized Sunday night. One possibility is that the Federal Reserve will support the move by opening its wholesale-borrowing window and relaxing collateral requirements for borrowers.

Lehman, a 158-year-old firm

Please turn to page A18

Ultimatum By Paulson Sparked Frenzied End

After more than a century and a half as an investment house, Lehman Brothers Holdings Inc. tottered on the brink Sunday night. But its latest troubles were set in motion days earlier.

In midweek, Treasury Secre-

By Deborah Solomon, Dennis K. Berman, Susanne Craig and Carrick Mollenkamp

Finance

United States

Housing

Economy

Multi Label Classification (MLC)

- MLC is the task of assigning a set of target labels for a given sample
- Given input ***x***, predict the set of labels $\{y_1, y_2, \ldots, y_L\}$, $y_i \in \{0,1\}$

x

y_1 **Castle**

y_2 **City**

y_3 **Mountains**

y_4 **Car**

y_5 **Road**

Multi Label Classification (MLC)

- MLC is the task of assigning a set of target labels for a given sample
- Given input ***x***, predict the set of labels $\{y_1, y_2, \ldots, y_L\}$, $y_i \in \{0,1\}$

x

y_1 **Castle** ✓

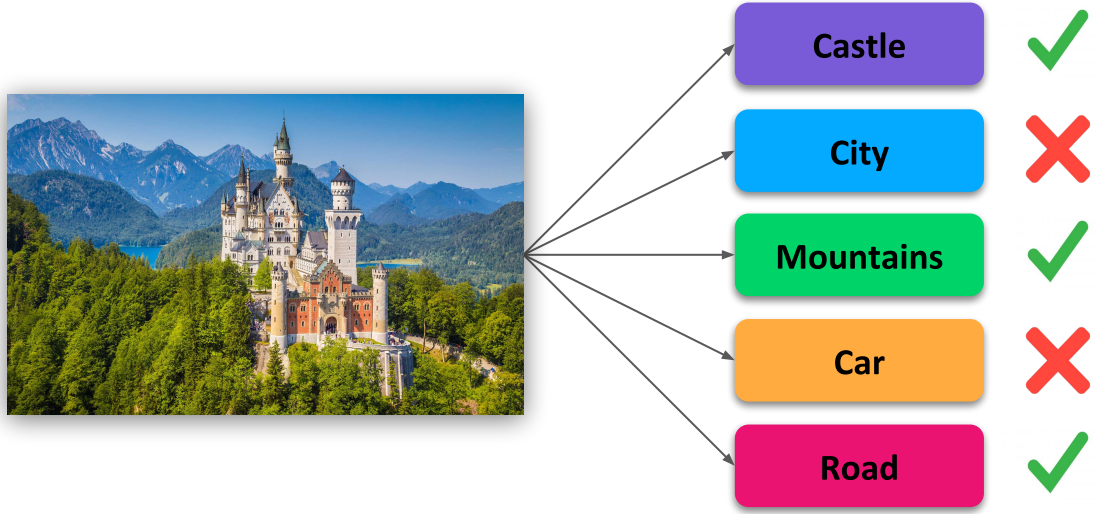
y_2 City ✗

y_3 **Mountains** ✓

y_4 Car ✗

y_5 **Road** ✓

Binary Relevance Classifiers



$$p(\mathbf{y}|\mathbf{x}) \approx \prod_{i=1}^{L} p(y_i|\mathbf{x})$$

Binary Relevance Classifiers

Castle

City

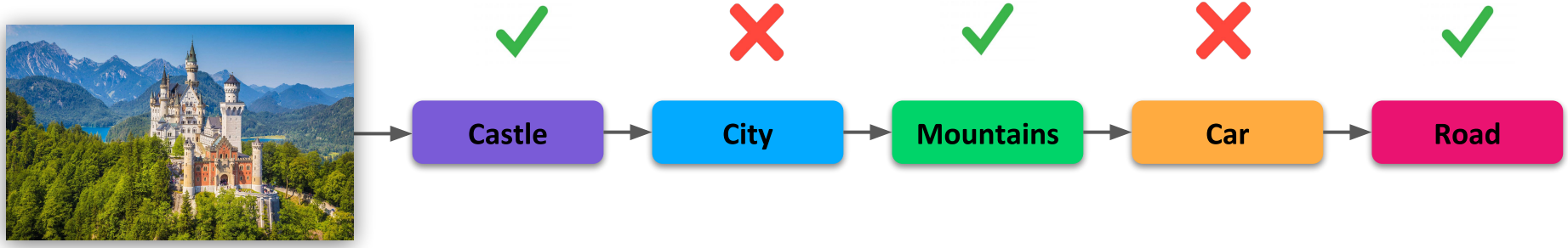
Road

There are dependencies between labels!

$$p(\mathbf{y}|\mathbf{x}) \approx \prod_{i=1}^{L} p(y_i|\mathbf{x})$$

Probabilistic Chain Classifiers

Read et. al. 2009, Wang et. al 2016, Nam et. al. 2017



$$p(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{L} p(y_i|y_{1:i-1}, \mathbf{x})$$

Deep Structured Prediction Classifiers

Belanger & McCallum 2016, Gygli et. al. 2017

Castle

City

Mountains

Car

Road

Energy Function E

$$\mathbf{y} = \underset{\bar{\boldsymbol{y}} \in \{0,1\}^L}{\arg\min}\, E\left(\bar{\boldsymbol{y}}; F(\boldsymbol{x})\right)$$

Castle ✓

City ✗

Mountains ✓

Car ✗

Road ✓

Embedding Models

Bhatia et. al. 2015, Wu et. al. 2017

Castle

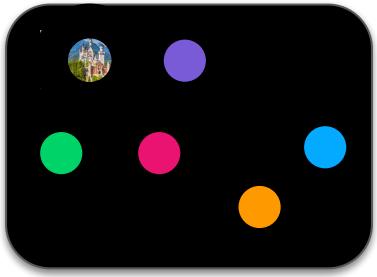
City

Mountains

Car

Road

Embedding Space



$\text{sim}(\mathbf{x}, y_1)$, $\text{sim}(\mathbf{x}, y_2)$, ...

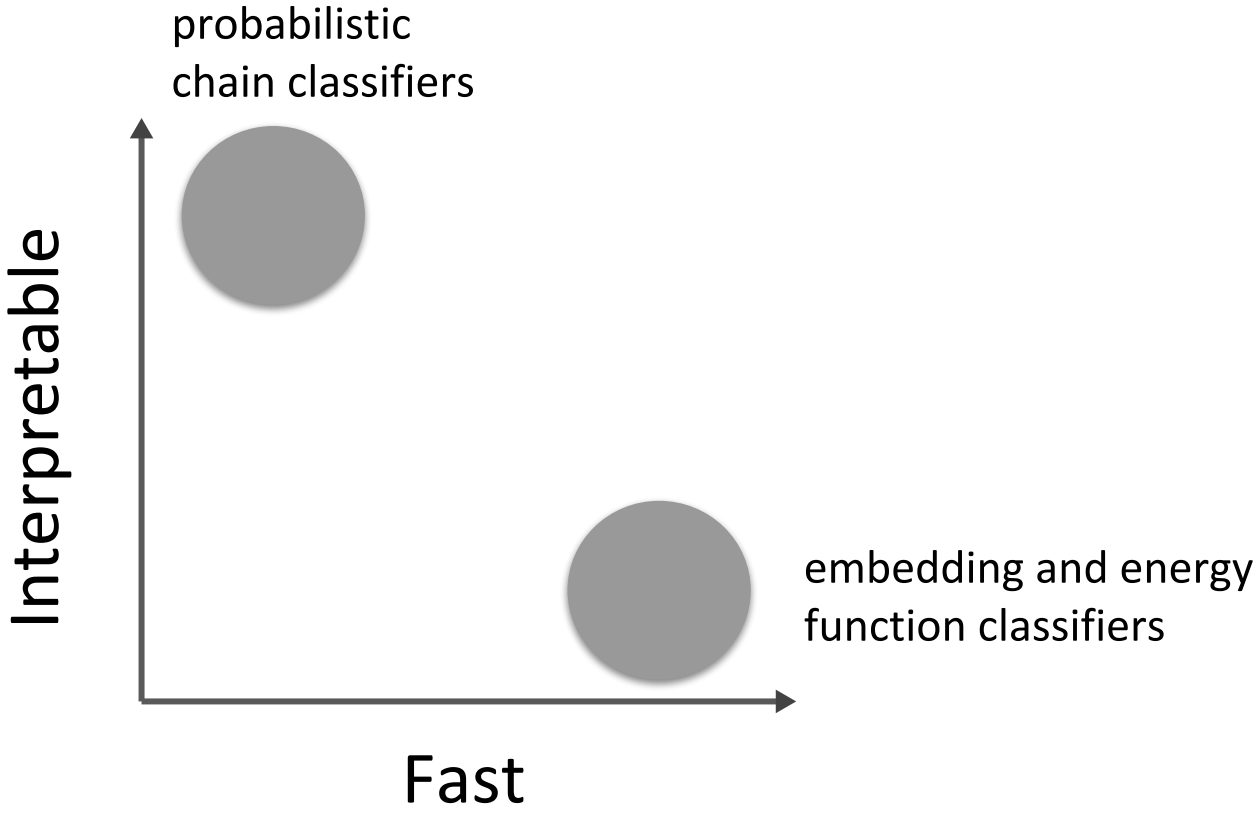
Castle ✓

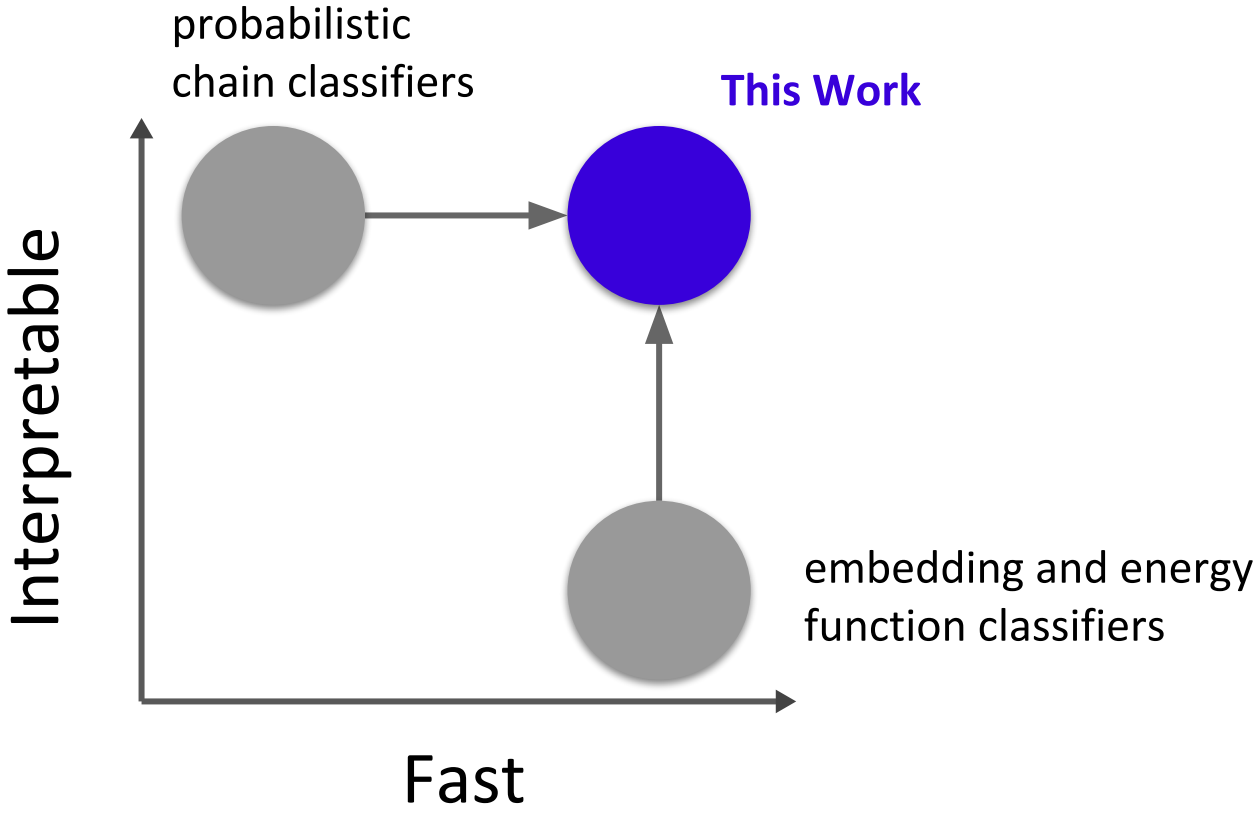
City ✗

Mountains ✓

Car ✗

Road ✓

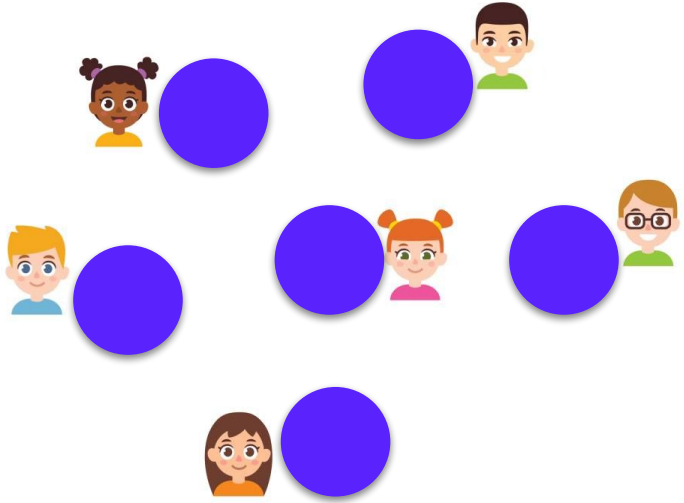

probabilistic
chain classifiers
Interpretable
embedding and energy
function classifiers
Fast


probabilistic
chain classifiers
This Work
Interpretable
embedding and energy
function classifiers
Fast

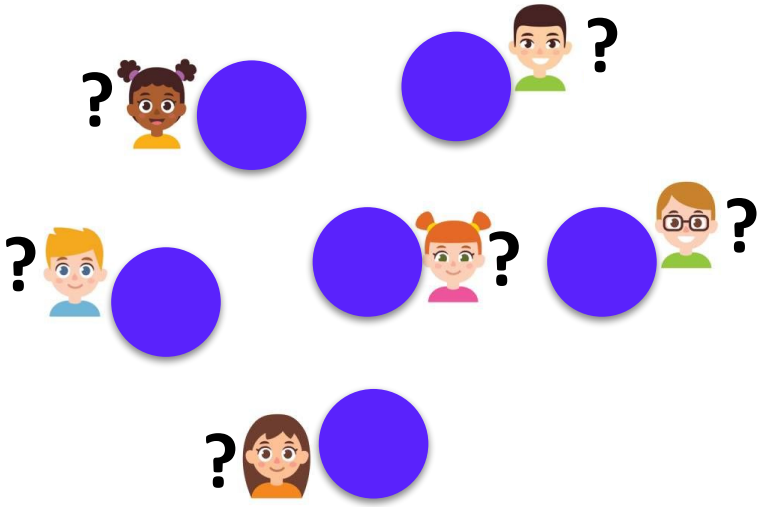
Background: Message Passing Neural Networks

Message Passing Neural Networks (MPNNs)

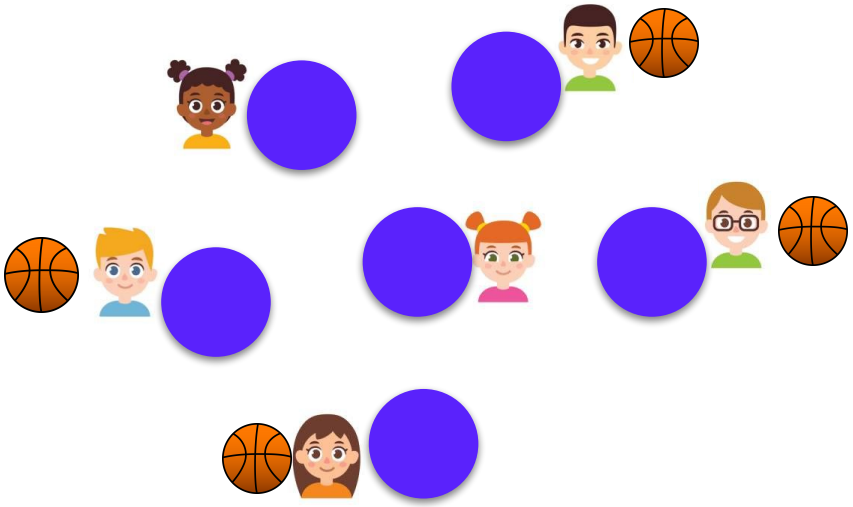
Generalization of Graph Neural Networks



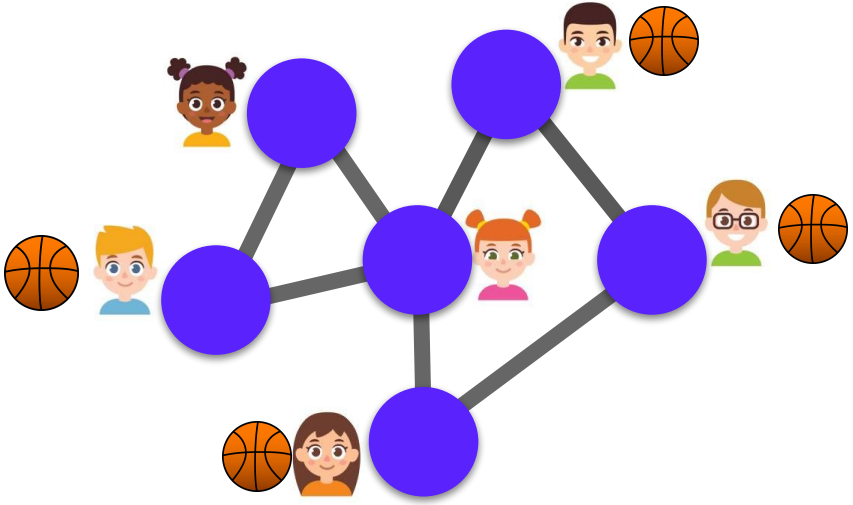
Message Passing Neural Networks (MPNNs)



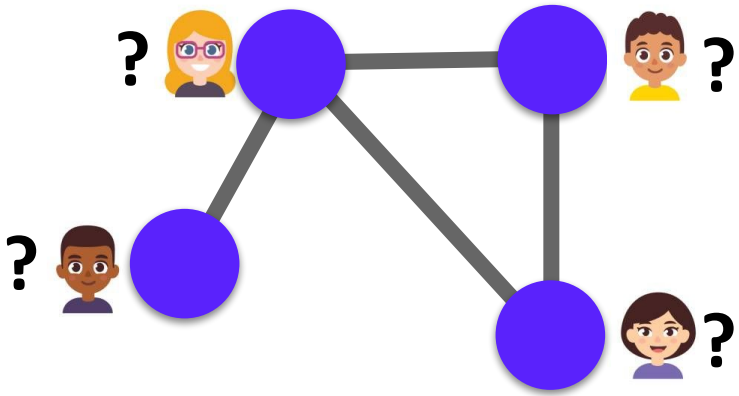
Message Passing Neural Networks (MPNNs)



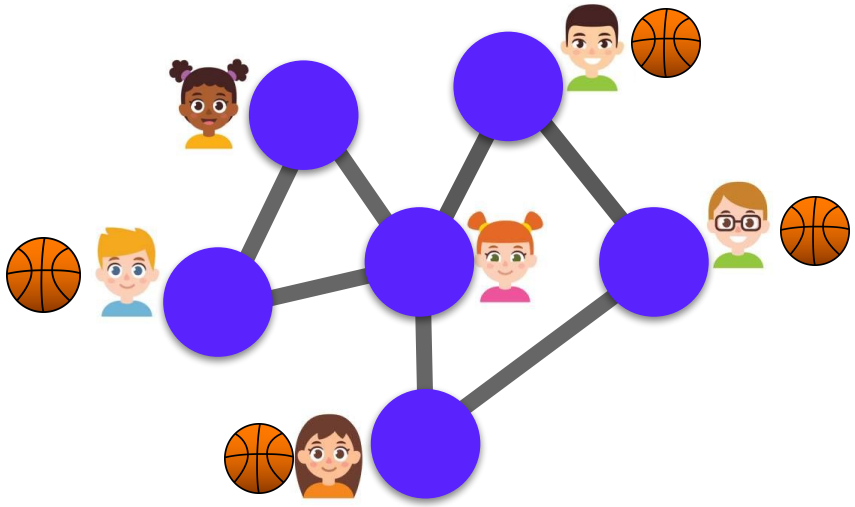
Message Passing Neural Networks (MPNNs)



Message Passing Neural Networks (MPNNs)

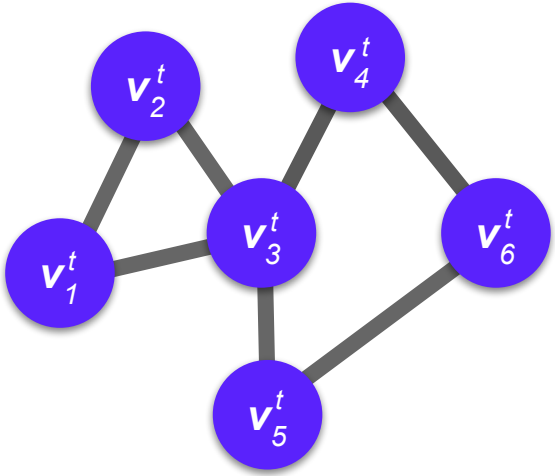


Message Passing Neural Networks (MPNNs)



- MPNNs can learn useful node representations for classifying nodes by encoding local graph structures and node attributes
- **Main idea:** pass messages between pairs of nodes and update them

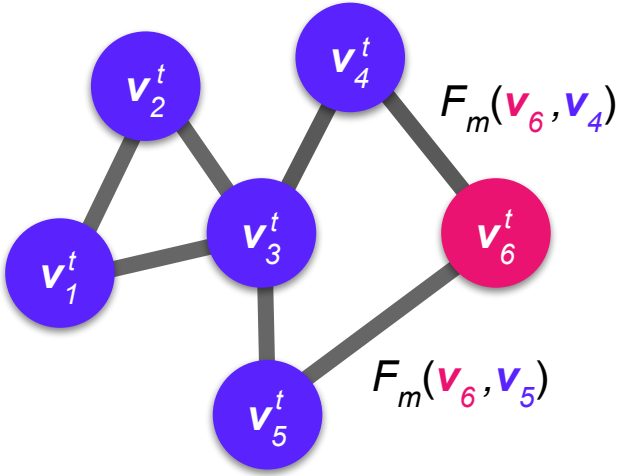
Message Passing Neural Networks (MPNNs)



$$G = (V, E)$$

$$\boldsymbol{v}_i^t \in \mathbb{R}^d$$

Message Passing Neural Networks (MPNNs)



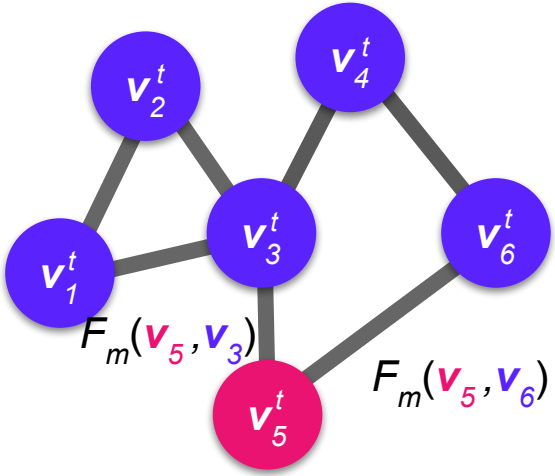
$$G = (V, E)$$

$$\boldsymbol{v}_i^t \in \mathbb{R}^d$$

message function F_m $\quad \boldsymbol{m}_i^t = \sum_{j \in \mathcal{N}(i)} F_m(\boldsymbol{v}_i^t, \boldsymbol{v}_j^t),$

node update function F_u $\quad \boldsymbol{v}_i^{t+1} = F_u(\boldsymbol{m}_i^t)$

Message Passing Neural Networks (MPNNs)



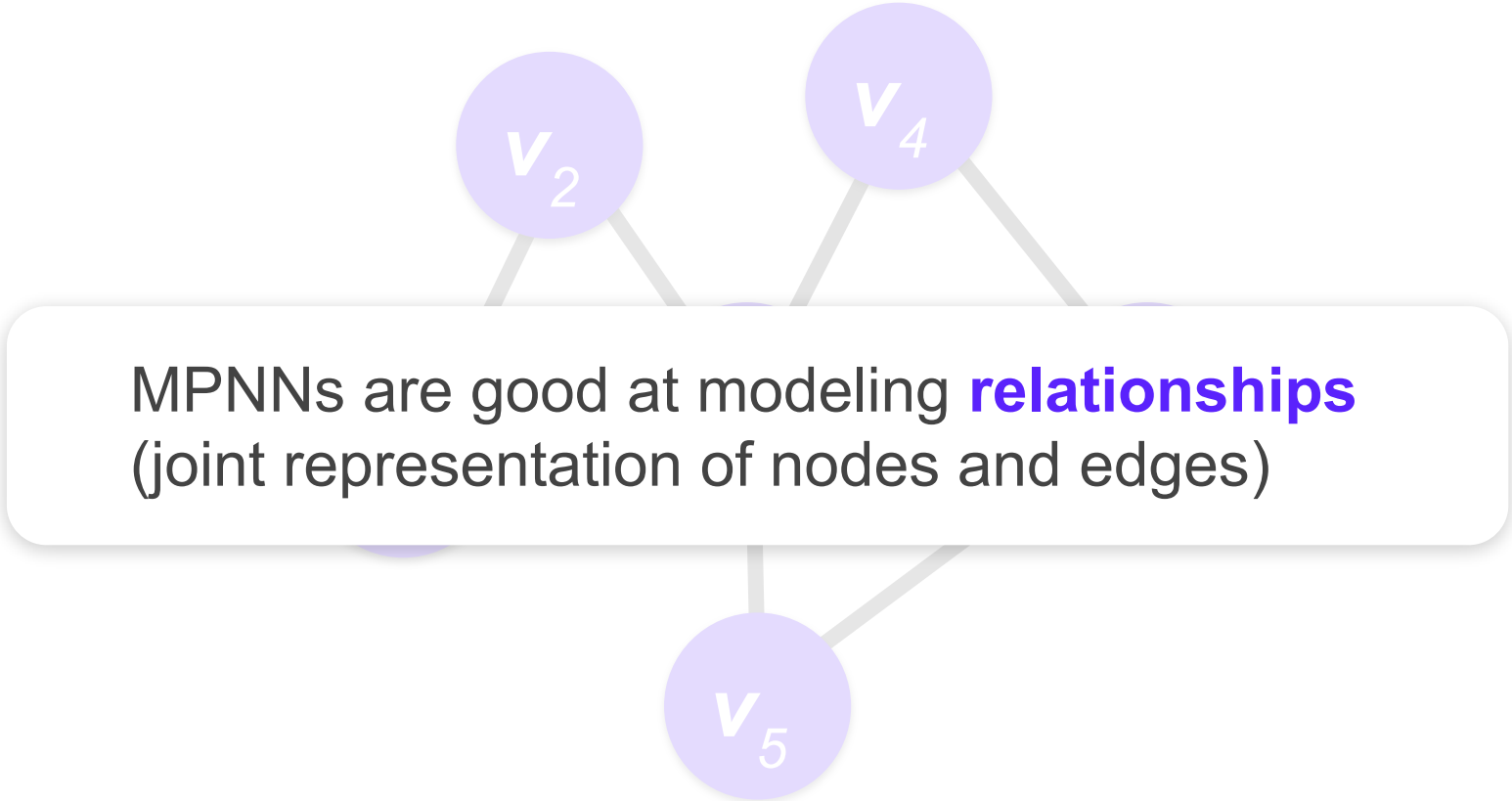
$$G = (V, E)$$

$$\boldsymbol{v}_i^t \in \mathbb{R}^d$$

message function F_m $\quad \boldsymbol{m}_i^t = \sum_{j \in \mathcal{N}(i)} F_m(\boldsymbol{v}_i^t, \boldsymbol{v}_j^t),$

node update function F_u $\quad \boldsymbol{v}_i^{t+1} = F_u(\boldsymbol{m}_i^t)$

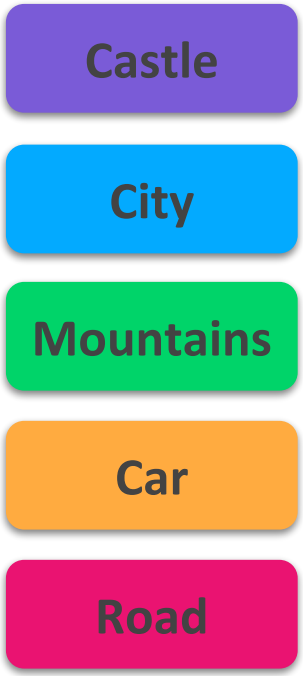
Message Passing Neural Networks (MPNNs)



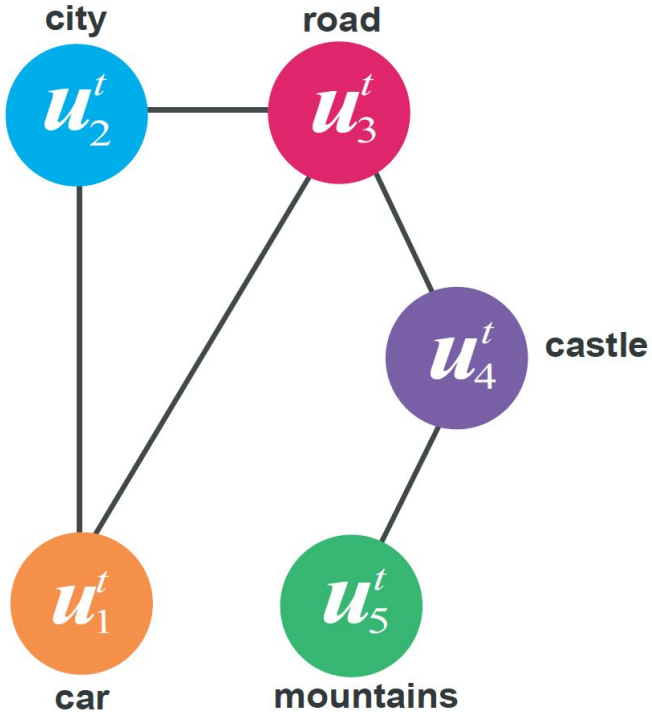
MPNNs are good at modeling **relationships** (joint representation of nodes and edges)

MPNNs for Multi Label Classification

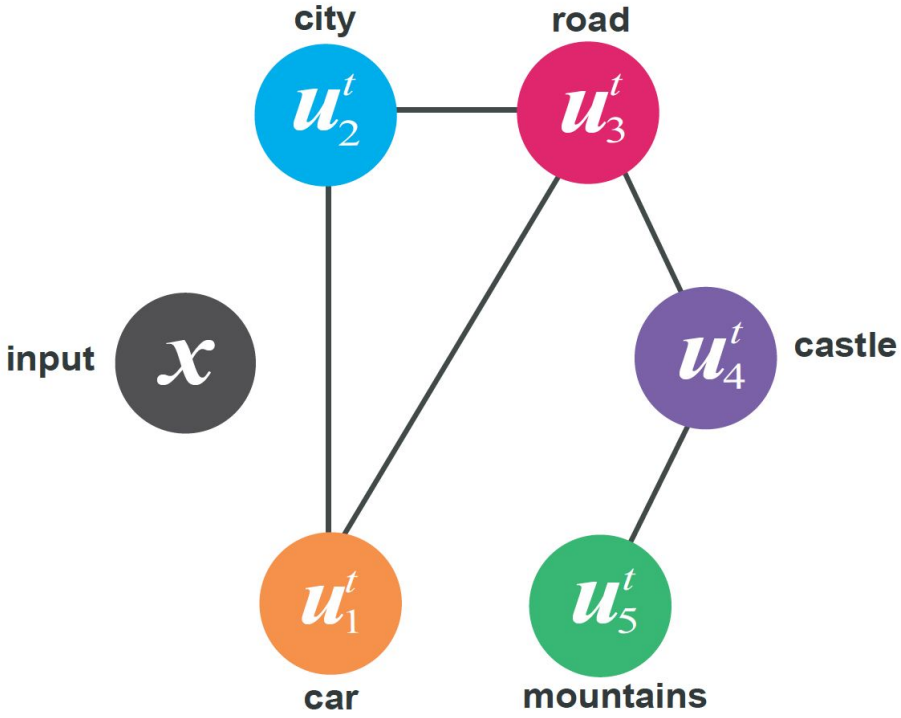
Neural Message Passing for MLC



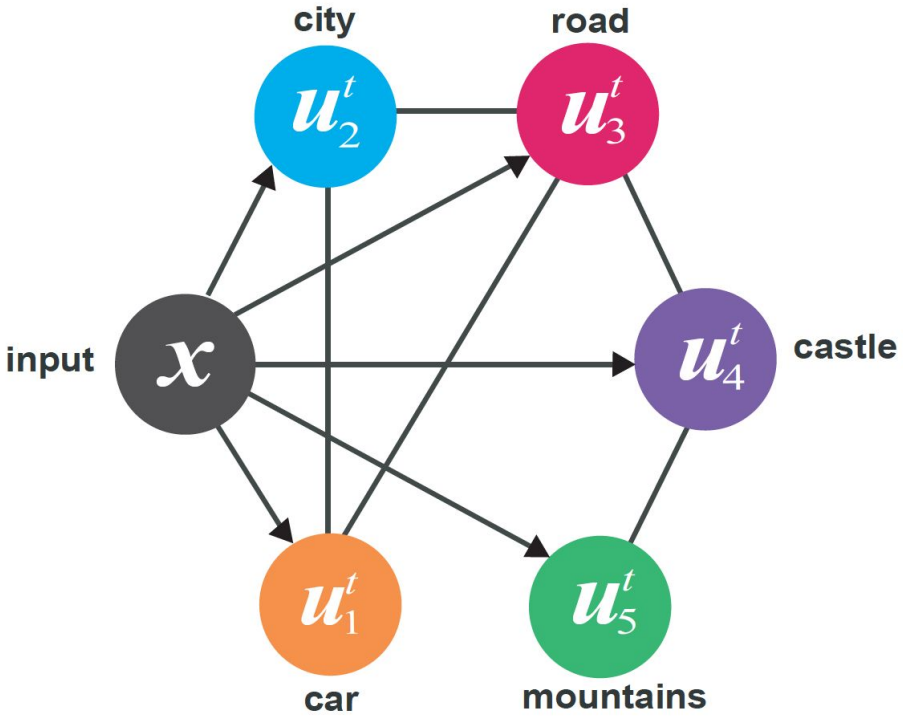
Neural Message Passing for MLC



Neural Message Passing for MLC

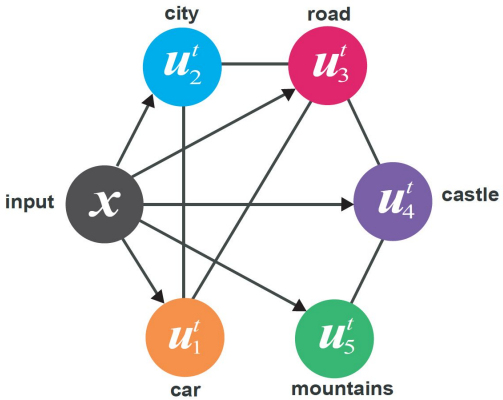


Neural Message Passing for MLC



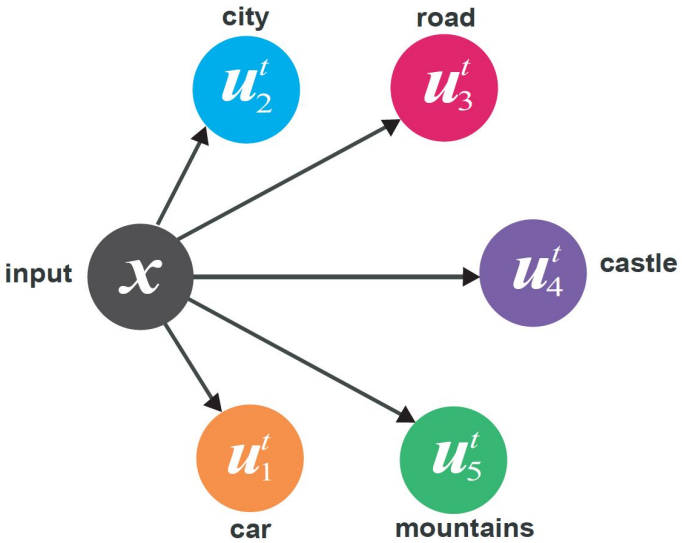
This Paper: Label Message Passing (LaMP) Networks

- **Main Idea:** Labels are represented as nodes in a label-interaction graph
- Given input embedding ***x***, the goal of LaMP is to model the conditional dependencies between label embeddings $\{\boldsymbol{u}^t_{1:L}\}$ using Message Passing Neural Networks
- LaMP uses MPNN modules to update label embeddings at each step *t* in two parts



Part A: Input-to-Label Message Passing

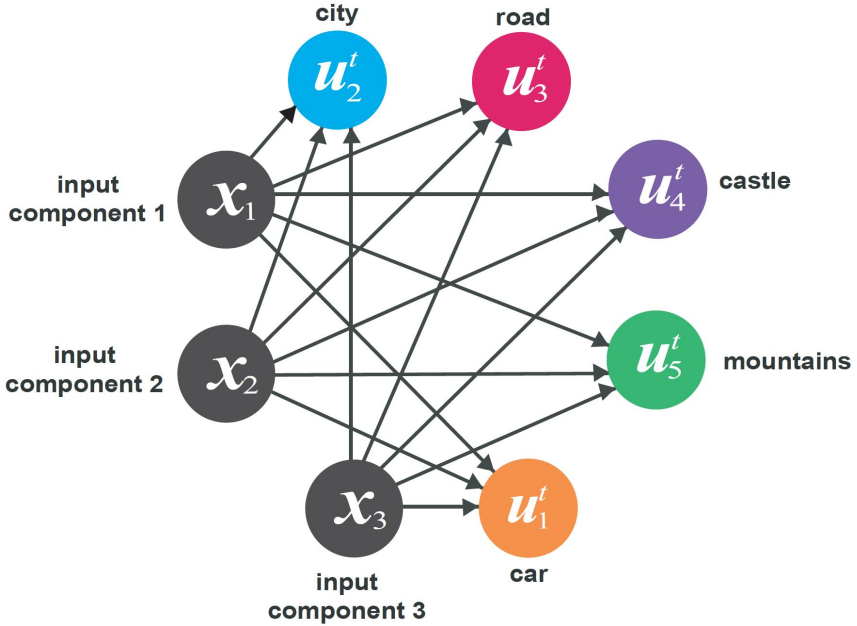
The first step passes messages from input $\boldsymbol{x}$ to each label embedding $\boldsymbol{u}_i^t$ to produce the intermediate embedding state $\boldsymbol{u}_i^{t'}$



$$\boldsymbol{m}_i^t = F_m(\boldsymbol{u}_i^t, \boldsymbol{x}),$$
$$\boldsymbol{u}_i^{t'} = F_u(\boldsymbol{m}_i^t).$$

Part A: Input-to-Label Message Passing

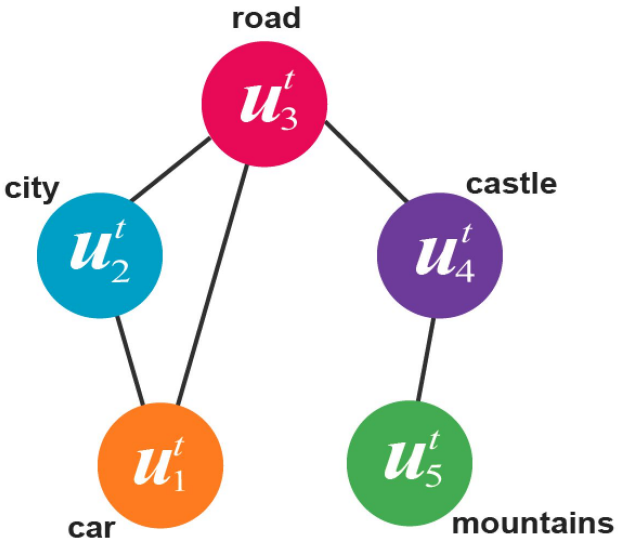
The first step passes messages from input $\boldsymbol{x}$ to each label embedding $\boldsymbol{u}_i^t$ to produce the intermediate embedding state $\boldsymbol{u}_i^{t'}$



$$\boldsymbol{m}_i^t = \Sigma_{j \in \boldsymbol{x}} F_m(\boldsymbol{u}_i^t, \boldsymbol{x}_i),$$
$$\boldsymbol{u}_i^{t'} = F_u(\boldsymbol{m}_i^t).$$

Part B: Label-to-Label Message Passing

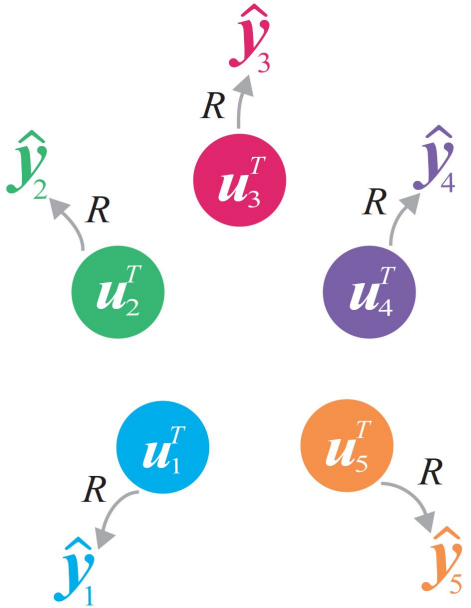
The second step passes messages between label embeddings to update their states conditioned on $\boldsymbol{x}$ to produce updated embedding $\boldsymbol{u}_i^{t+1}$:



$$\boldsymbol{m}_i^{t'} = \Sigma_{j \in \mathcal{N}(i)} F_m(\boldsymbol{u}_i^{t'}, \boldsymbol{u}_j^{t'}),$$
$$\boldsymbol{u}_i^{t+1} = F_u(\boldsymbol{m}_i^{t'}).$$

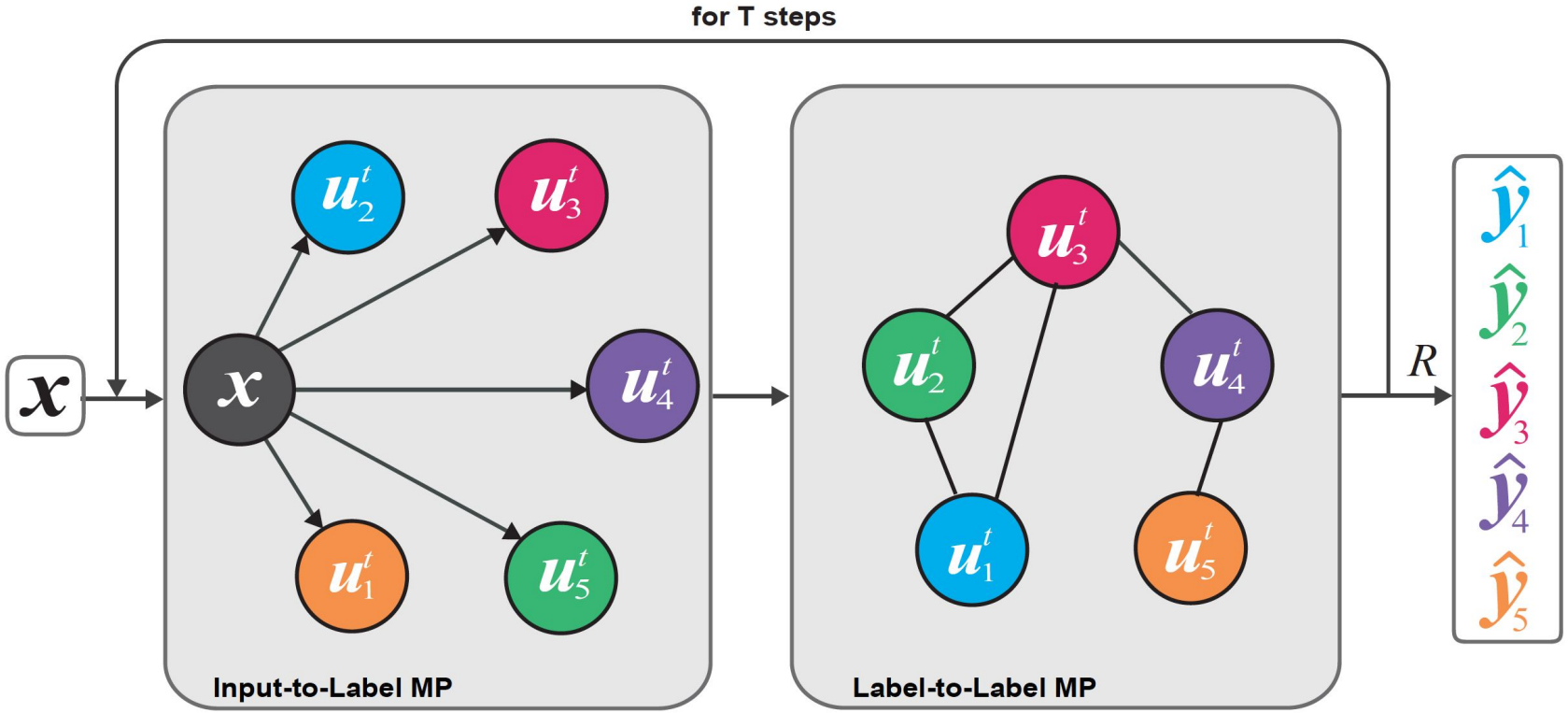
Readout Function

- After T updates to the label embeddings, the last module predicts the probabilities of each label being positive $\{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_L\}$
- Readout function R projects each of the L label embeddings $\{\boldsymbol{u}_{1:L}^T\}$ into a scalar value



$$\hat{y}_i = R(\boldsymbol{u}_i^T; \mathbf{W}^o) = \text{sigmoid}(\mathbf{W}_i^o \boldsymbol{u}_i^T)$$

Label Message Passing (LaMP) Networks



Loss Function

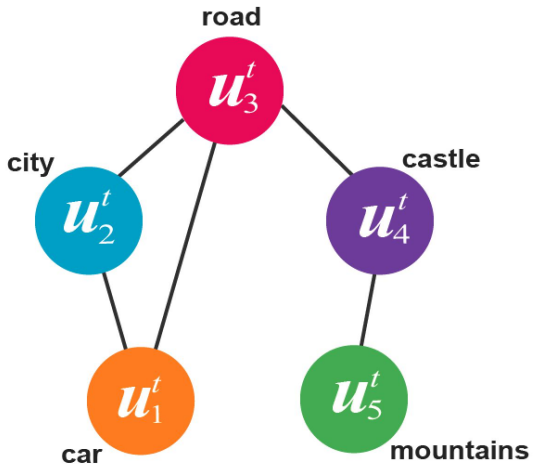
Binary cross entropy, however since LaMP iteratively updates the label state for T steps, we can impose a loss at each step t by using the readout function to obtain intermediate predictions $\hat{y}_i^t = R(\boldsymbol{u}_i^t; \mathbf{W}^o)$

$$Loss(\boldsymbol{y}, \hat{\boldsymbol{y}}^t) = \frac{1}{T} \sum_{t=0}^{T} \frac{1}{L} \sum_{i=1}^{L} -(y_i \log(\hat{y}_i^t) + (1 - y_i) \log(1 - \hat{y}_i^t))$$

Label Graph Structure

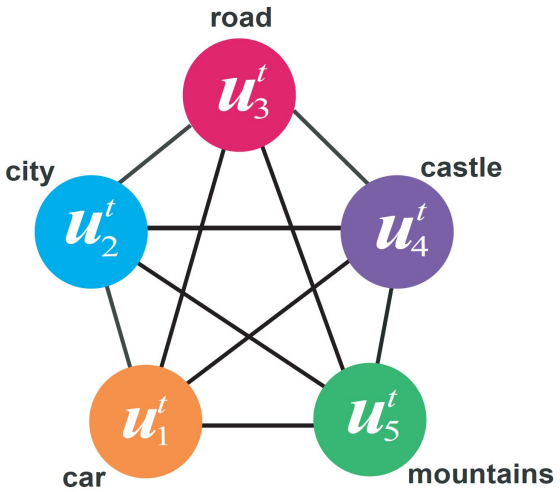
Prior Label Graph Structure

- For datasets with known label structure (hierarchy, protein interaction), use known
- For all other datasets, we place an edge on the adjacency matrix for all labels that co-occur in any sample of the training set

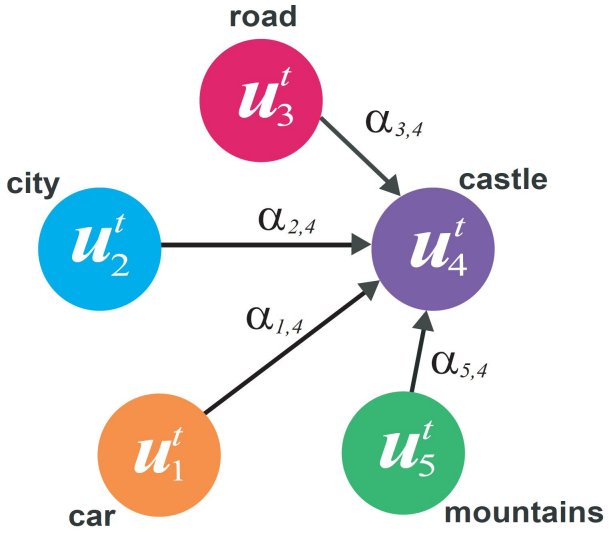


Learned Label Graph Structure Using Attention

- Prior graphs may not be the best assumption since they are not conditioned on inputs
- We propose using attention to learn the graph end-to-end while training the classifier



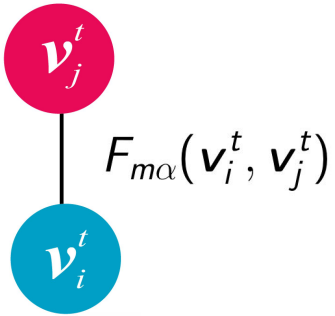
Attention-Based MPNNs



$$m_i^t = \Sigma_{j \in \mathcal{N}(i)} F_{m\alpha}(v_i^t, v_j^t),$$

$$F_{m\alpha}(v_i^t, v_j^t) = \alpha_{ij}^t \mathbf{W}^v v_j^t$$

Attention-Based MPNNs



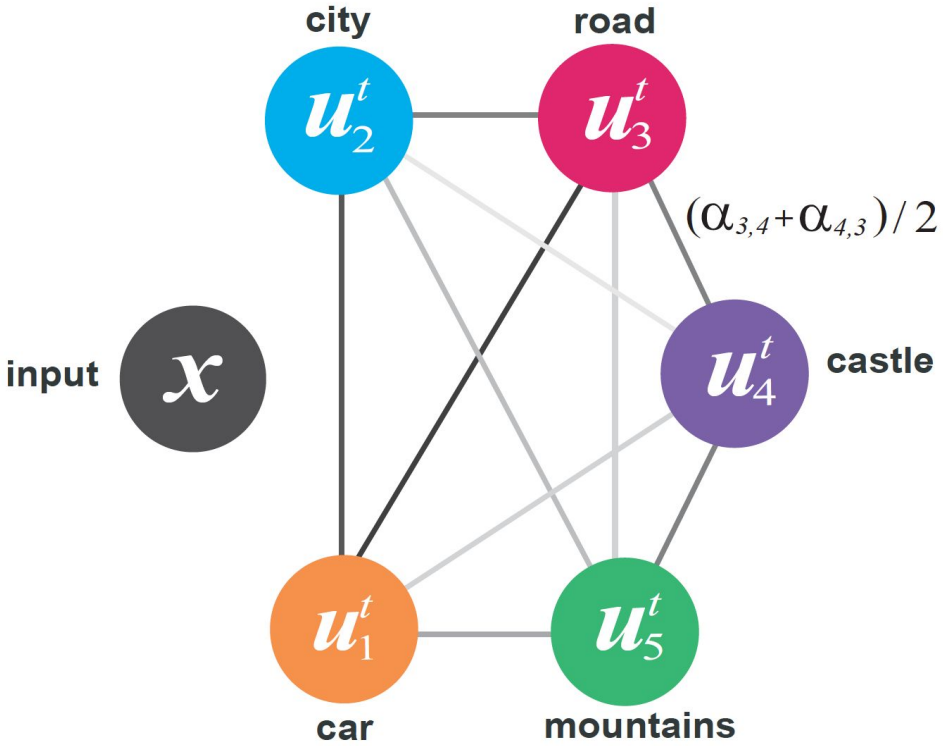
Attention coefficient e_{ij}^t for pair of nodes $(\boldsymbol{v}_i^t, \boldsymbol{v}_j^t)$ is computed using some attention function producing a scalar value representing the relationship between nodes:

$$e_{ij}^t = (\mathbf{W}^q \boldsymbol{v}_i^t)^\top (\mathbf{W}^u \boldsymbol{v}_j^t)$$

Attention weight α_{ij}^t produced by normalizing over neighboring nodes

$$\alpha_{ij}^t = \mathrm{softmax}_j(e_{ij}^t) = \frac{\exp(e_{ij}^t)}{\sum_{k \in \mathcal{N}(i)} \exp(e_{ik}^t)}.$$

Attention-Based MPNNs



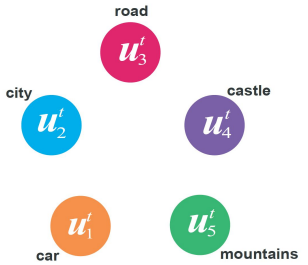
Experiments and Results

Datasets

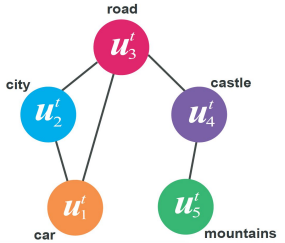
- We validate our model on 8 real world MLC datasets
- These datasets vary in the number of samples, number of labels, input type (sequential, tabular, graph, vector), and output type (unknown label graph, prior label graph)

LaMP Variations

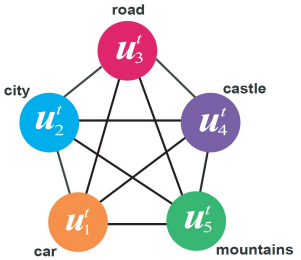
LaMP$_{el}$ uses an edgeless label graph and messages are not passed between labels, assuming no label dependencies



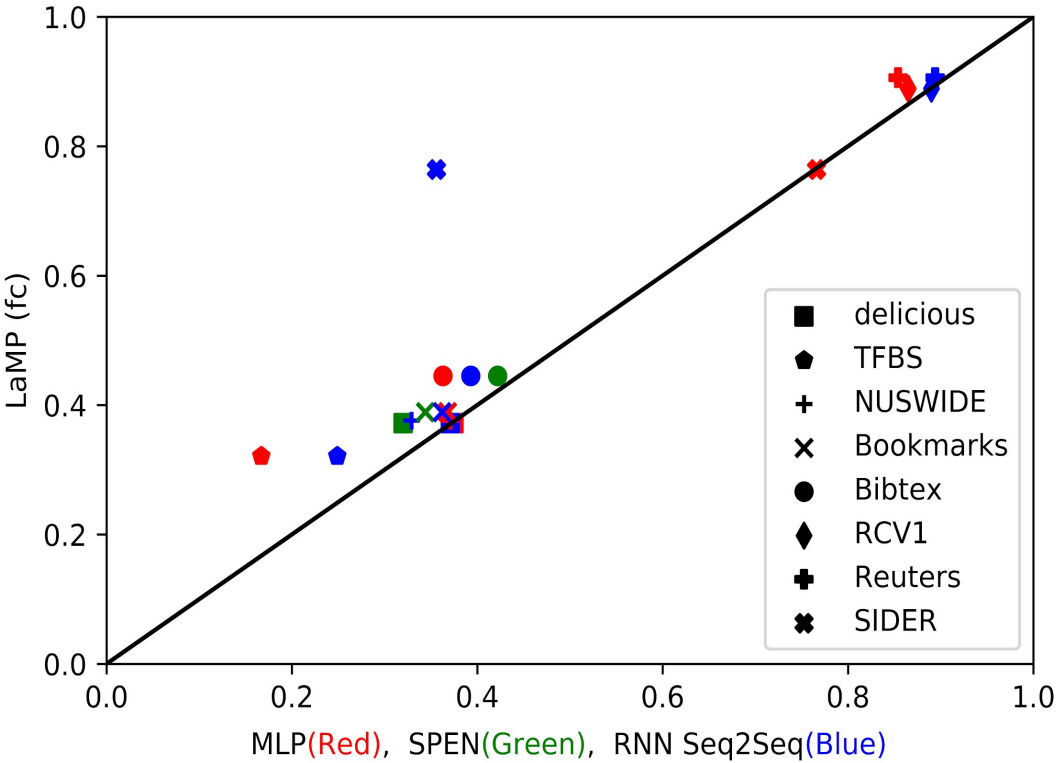
LaMP$_{pr}$ uses a prior label graph where each label is able to attend to only other labels from the known label graph



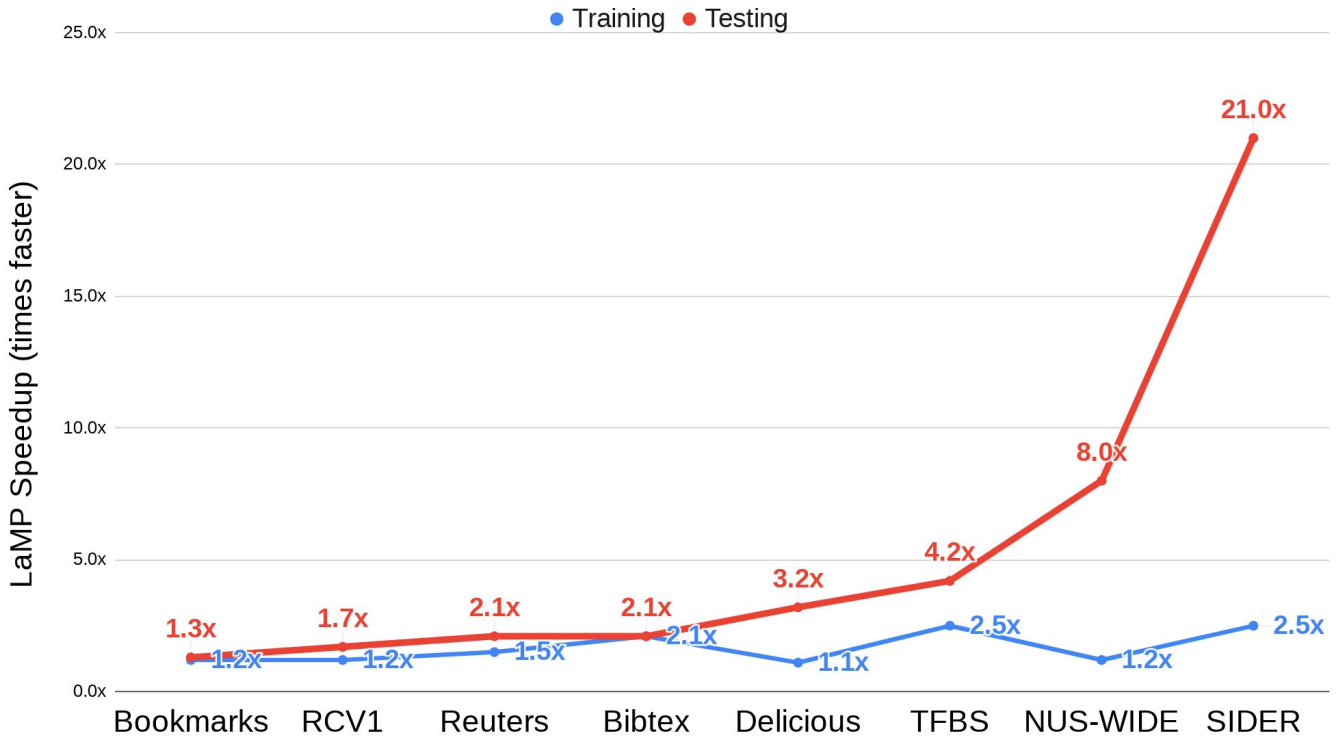
LaMP$_{fc}$ uses a fully connected label graph where each label is able to attend to all other labels



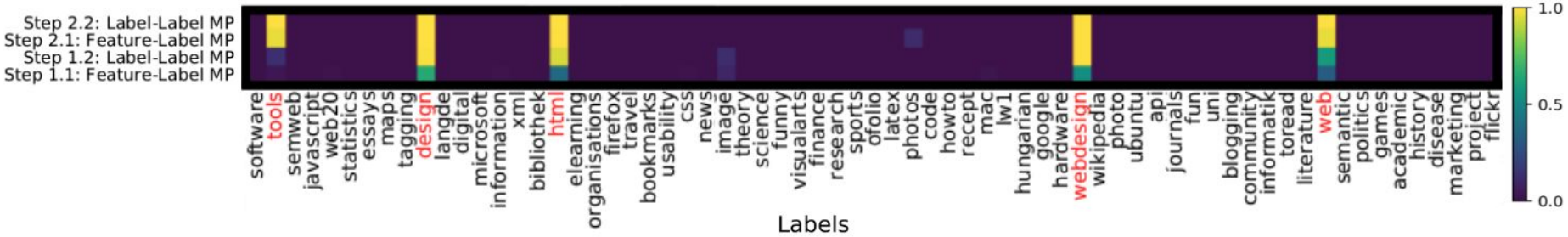
Example-Based F1 Scores (ebF1)



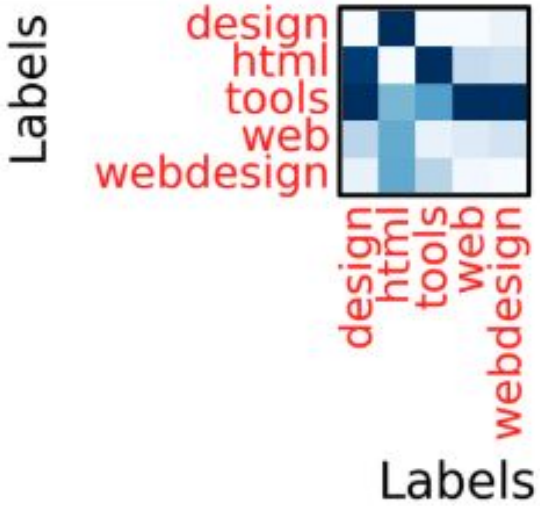
LaMP vs RNN Speed



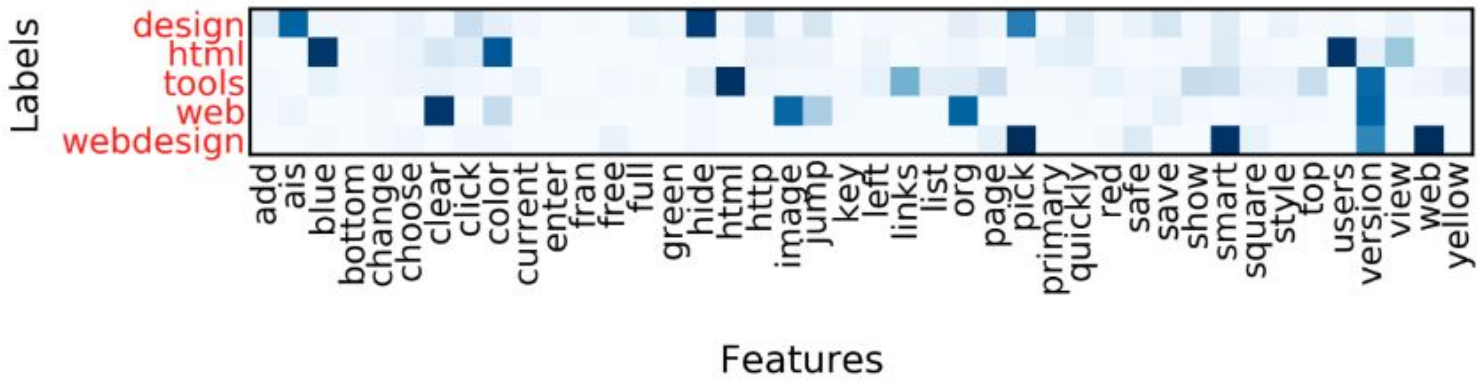
Visualization: Intermediate Predictions



Visualization: Label to Label Attention

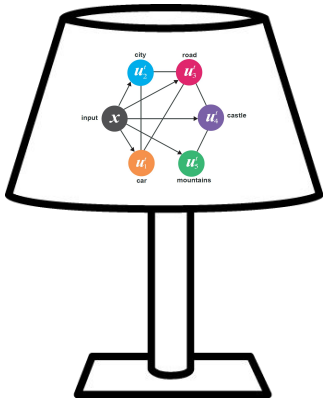


Visualization: Label to Input Attention



Conclusion

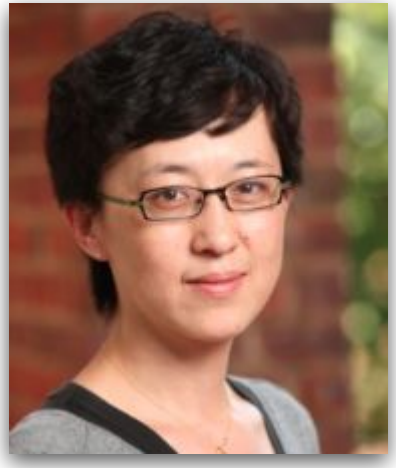
- LaMP models label interactions for MLC by **placing labels as nodes on a graph**
- LaMP networks are as **accurate, faster, and more interpretable** than the previous state-of-the-art MLC classifiers
- **Representation learning**: exploiting the inductive bias of a model automatically discover the representations needed for classification



Thank You

code available at: github.com/QData/LaMP

Arshdeep Sekhon



Dr. Yanjun Qi

UNIVERSITY of VIRGINIA