CENTRE FOR ARTIFICIAL INTELLIGENCE RESEARCH

Tsetlin Machine Applications

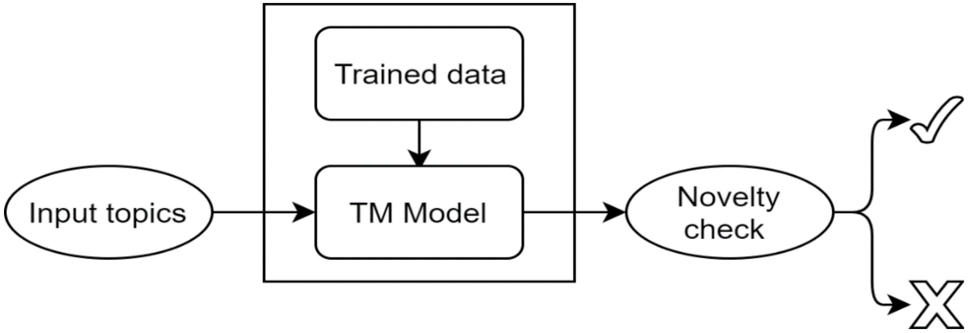
BIMAL BHATTARAI

University of Agder

Novelty in NLP

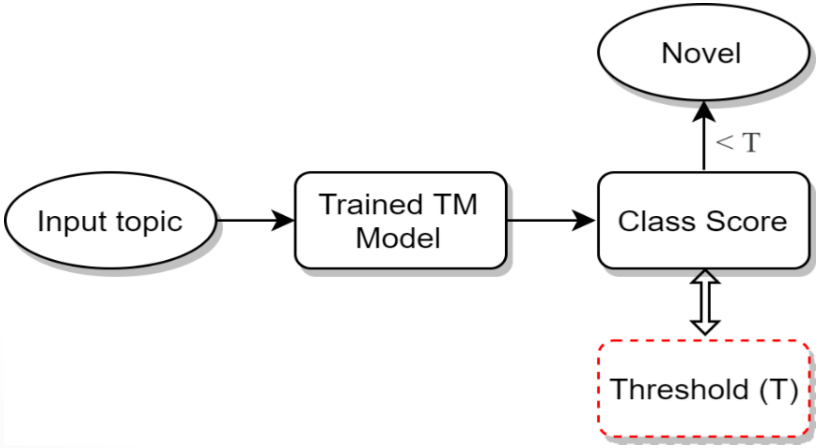
Novelty Detection

Determine if the given input topic is novel or not



Novelty Description

Describe the novelty of the input topic based on the clauses



Example: Novelty Detection

Trained topics: Account and Card queries
Users Queries: Loan (New topic)

Example:
Intent – I have an account in this bank, How can I get a loan?
Answer- (can be anything)

Tsetlin Working:

Example:
Intent – I have an account in this bank, How can I get a loan?
TM Model- *IF* {(account) *and* (bank) *and* (get) *and* (loan)} *then* New topic
Novelty Score- Very low

Example: Novelty Description

Sentence: The apple that he ate yesterday was good**.**

Category- Novel

Sentence : The apple shares decreased over this week.

Category- Not Novel (Known)

- Different scores for “apple”: high for “apple” fruit, low for “apple” company/phone

Results

Novelty Detection

Algorithms	20 Newsgroup	Spooky action author	CMU movie	BBC sports	WOS
LOF	65.3 %	47.94 %	56.68 %	67.88 %	71.08 %
Feature Bagging	65.40 %	48.83 %	57.42 %	64.23 %	69.64 %
HBOS	73.90 %	27.08 %	46.26 %	32.74 %	**86.01 %**
Isolation Forest	78.50 %	36.18 %	53.36 %	35.93 %	72.43 %
Average KNN	74.30 %	43.61 %	56.78 %	65.69 %	67.98 %
K-Means clustering	81.00 %	61.30 %	49.20 %	47.70 %	41.31 %
One-class SVM	**84.30 %**	61.50%	67.15 %	92.74 %	77.17 %
TM framework	83.00 %	**63.68 %**	**71.86 %**	**93.10 %**	76.08 %

Table 2: Performance comparison of proposed TM framework with cluster and outlier based novelty detection algorithms.

Novelty Description

Table 2 Relative frequency and score for each word

Known				Novel			
Word	Frequency	Relative frequency	Score	Word	Frequency	Relative frequency	Score
England	1	0.071	1.070	England	1	0.076	1.070
Won	1	0.071	2.169	Won	2	0.154	2.169
Cricket	4	0.28	0.271	Rugby	4	0.307	4.651
Match	1	0.071	2.169	Match	2	0.154	2.169
Hit	2	0.142	0.535	Despite	1	0.076	1.15
Six	4	0.28	0.271	Old	2	0.153	2.31
Ball	1	0.071	1.070	Ball	1	0.076	1.070

Problem: Data Representation

- Problem of Polysemy:

  Sentence 1: *He sent me a* ***present*** *for my birthday.*

  Sentence 2: *There were 200 people* ***present*** *at the meeting.*

- Problem of contextual dilemma:

  Sentence : *I am a _____, and I am in a class.*

  A) Student

  B) Teacher

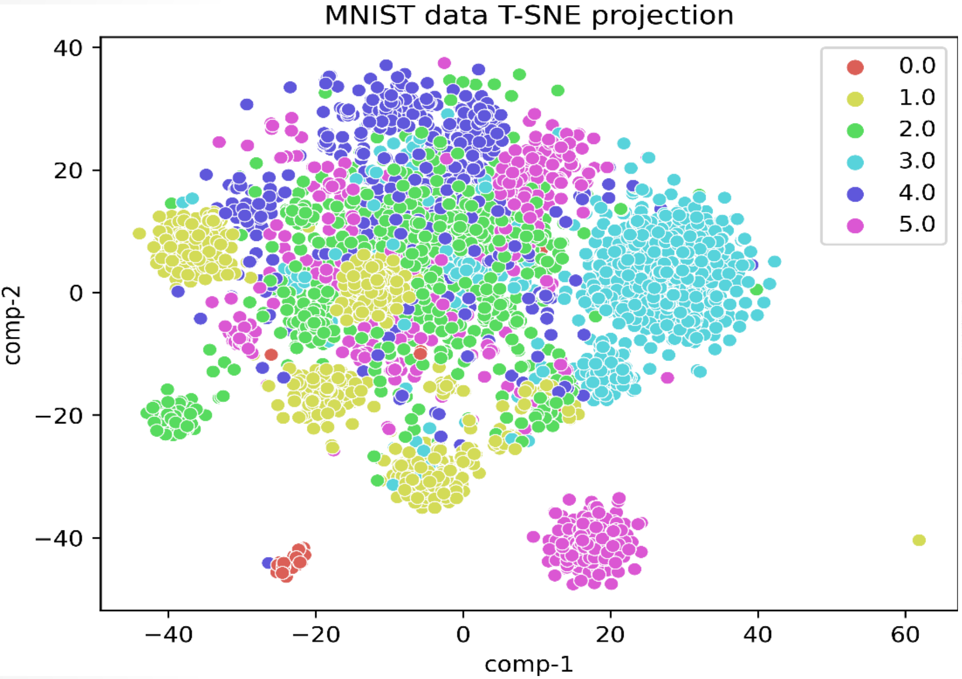
TM Representation

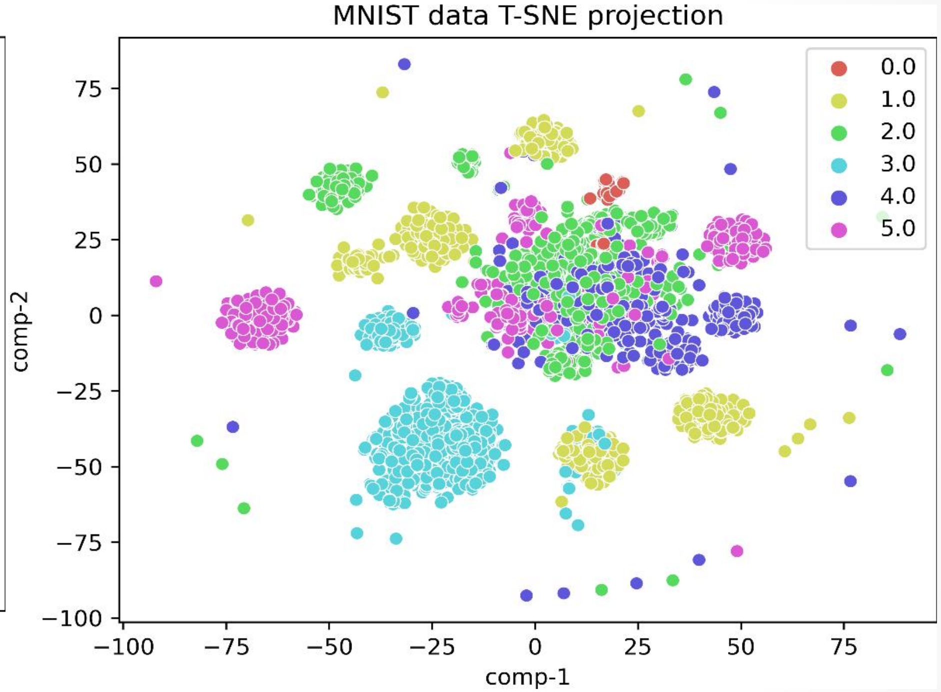
Example:
Feature – dog
Deep learning representation (embedding): *[0.35, 0.86, -0.36, …….., -0.21]*
TM representation: [*canine* ***and*** *loyal* ***and*** *domestic* **…… *and*** *bark***]**
TM Feature space: Binary (i.e., [1, 0, 1, ……, 1])

Practicality:
Class : Cancer
Deep learning representation: Embedding
If Noise is present, Prediction can be False

TM representation: Clauses Captured Features
Noise tolerant, Robust Prediction

Visualization





Result

Models	TREC	MPQA	SUBJ	WebKB	CR	R8
LSTM	87.19	89.43	85.66	85.32	80.06	96.09
BiLSTM	91.0	89.5	92.3	-	-	96.31
CNN-non-static	93.6	89.05	93.4	-	84.3	95.71
CNN-static	92.0	89.06	93.0	-	84.7	94.02
CNN-multichannel	92.2	89.4	93.2	-	85.0	-
DiSAN	94.2	90.1	94.2	-	84.8	-
BERT	**97.6**	**90.66**	**97.0**	79.0	**86.58**	96.02
TM	91.6	74.55	86.8	91.69	80.55	95.93
TM_{rep}	95.6	87.3	90.1	**93.05**	83.06	**96.84**

Table 2: Performance comparison of our model with baseline algorithms. We reproduce the results with the same hyperparameter configurations for all baselines for a fair comparison and report average accuracy across 10 different random seeds.

TABLE III
DOMAIN ADAPTATION PERFORMANCE (ACCURACY %) ON AMAZON REVIEW DATASET.

	S-only	MMD	DANN	CORAL	WDGRL	ACAN	BERT	TM_{rep}
$B \rightarrow D$	81.09	82.57	82.07	82.74	83.05	83.45	**86.75**	84.94
$B \rightarrow E$	75.23	80.95	78.98	82.93	83.28	81.20	82.80	**86.21**
$B \rightarrow K$	77.78	83.55	82.76	84.81	85.45	83.05	86.20	**87.57**
$D \rightarrow B$	76.46	79.93	79.35	80.81	80.72	82.35	81.55	**85.06**
$D \rightarrow E$	76.24	82.59	81.64	83.49	83.58	82.80	80.60	**86.81**
$D \rightarrow K$	79.68	84.15	83.41	85.35	86.24	78.60	83.00	**87.75**
$E \rightarrow B$	73.37	75.72	75.95	76.91	77.22	79.75	81.85	**84.83**
$E \rightarrow D$	73.79	77.69	77.58	78.08	78.28	81.75	**83.85**	83.43
$E \rightarrow K$	86.64	87.37	86.63	87.87	88.16	83.35	**90.80**	87.88
$K \rightarrow B$	72.12	75.83	75.81	76.95	77.16	80.80	82.10	**82.30**
$K \rightarrow D$	75.79	78.05	78.53	79.11	79.89	82.10	82.05	**83.07**
$K \rightarrow E$	85.92	86.27	86.11	86.83	86.29	86.60	**88.35**	88.31
AVG	77.84	81.22	80.74	82.16	82.43	82.15	84.13	**85.68**

TM Explainability

PolitiFact							
True				Fake			
Plain	times	Negated	times	Plain	times	Negated	times
trump	297	candidate	529	congress	136	trump	1252
said	290	debate	413	tax	104	profession	1226
comment	112	civil	410	support	70	navigate	1223
donald	110	reform	369	senate	64	hackings	1218
story	78	congress	365	president	60	reported	1216
medium	63	iraq	361	economic	57	arrest	1222
president	48	lawsuit	351	americans	49	camps	1206
reported	45	secretary	348	candidate	48	investigation	1159
investigation	38	tax	332	debate	44	medium	1152
domain	34	economy	321	federal	41	domain	1153

Table 3: Top ten Literals captured by clauses of TM for PolitiFact.

GossipCop							
True				Fake			
Plain	times	Negated	times	Plain	times	Negated	times
source	357	stream	794	season	150	insider	918
insider	152	aggregate	767	show	103	source	802
rumors	86	bold	723	series	79	hollywood	802
hollywood	80	refreshing	722	like	78	radar	646
gossip	49	castmates	721	feature	70	cop	588
relationship	37	judgment	720	video	44	publication	579
claim	33	prank	719	said	33	exclusively	551
split	32	poised	718	sexual	32	rumor	537
radar	32	resilient	714	notification	25	recalls	535
magazine	30	predicted	714	character	25	kardashian	525

Table 4: Top ten Literals captured by clauses of TM for GossipCop.

CAIR CENTRE FOR ARTIFICIAL INTELLIGENCE RESEARCH

Thank you!