Prairie View A&M University, TX, 77446

Presentation on:

Fake news detection using Temporal Ensembling model

Name: Uboho Victor T.

ID: P21443445

Course: Statistical Learning for Big Data

Instructor: Dr. Lijun Qian

Special thanks: Dr. Xishuang Dong

Outline

- Introduction
- Motivation
- Temporal Ensembling Model
- Data Preprocessing
- Experimental Results
- Future work
- Conclusion
- Question and answers

Introduction

• What is semi-supervised learning

- Semi-supervised machine learning falls in between the supervised and unsupervised learning
- In semi-supervised learning, an algorithm learns from a dataset that includes both labeled and unlabeled data, usually mostly unlabeled.

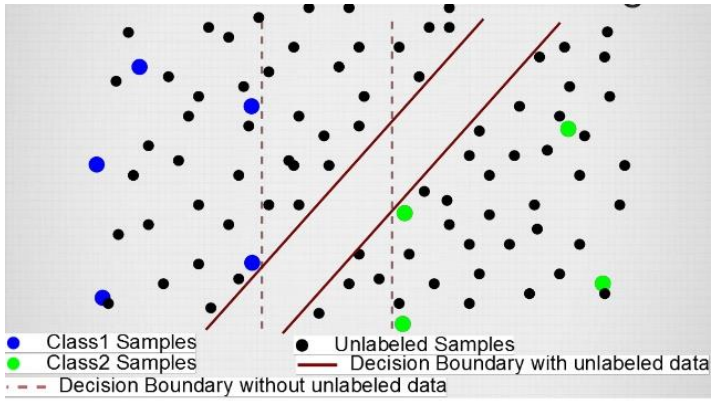


Fig 1: Semi-supervised learning

• Why semi-supervised learning?

Semi-supervised learning takes advantage of a huge amount of unclassified data, to perform a classification in situations when, typically, there is little labeled data which is a representation of where big data is today.

Motivation

- What is Fake news?

It describe factually incorrect and misleading articles published mostly for the purpose of making money through page views.

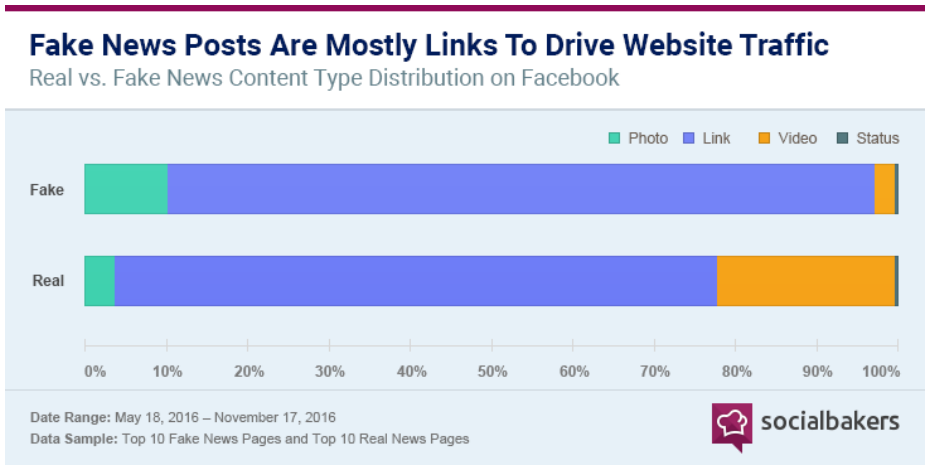


Chart 1: Real vs. Fake news content type distribution on Facebook - https://www.socialbakers.com

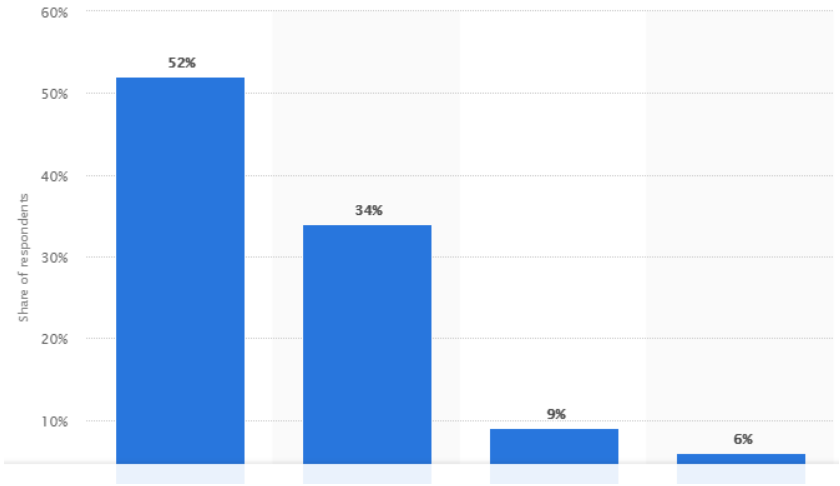


Chart 2: Perceived frequency of online news websites reporting fake news stories in the United States as of March 2018 - https://www.statista.com

- Fake news detection

- Understanding on what Fake News is.
- Techniques in the fields of machine learning, natural language processing helps us to detect fake news.

Motivation

- Semi-supervised learning for fake news detection

- *Fake news data is rapidly growing and new words are continuously generated*
- *Insufficient labeled data*
- *It helps to label new data*

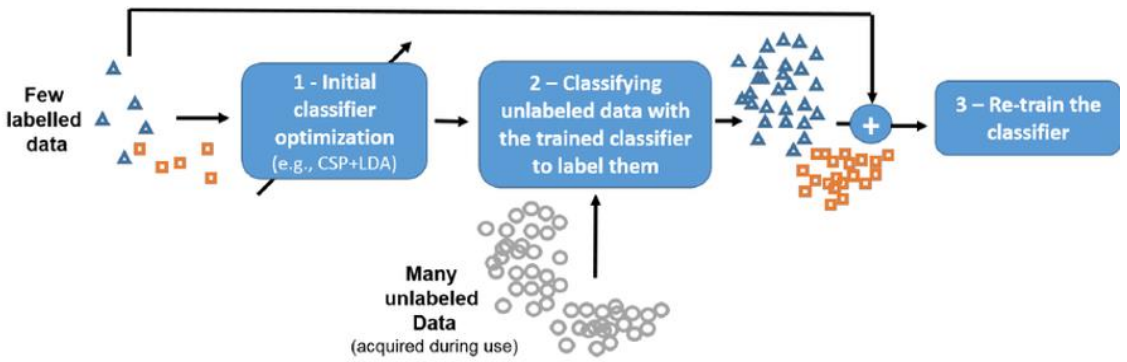


Fig 2: How semi-supervised learning helps to label new data

- Applications

- Government and Politics (Voting and elections)
- Business and Economy (Stock exchange), etc.

Our model for Fake news detection:

Temporal Ensembling model

Temporal Ensembling Model

- Pi Model

- Temporal Ensembling is a variance of the Π-MODEL
- *Encourages consistent network output between two realizations of the same input stimulus, under two different dropout conditions*
  - *Evaluate the network for each training twice*
  - *Loss Function is the sum of two components:*
    - *Component 1. Standard cross entropy loss for labelled*
    - *Component 2. Weighted Mean squared difference of the two prediction vectors (for all inputs)*

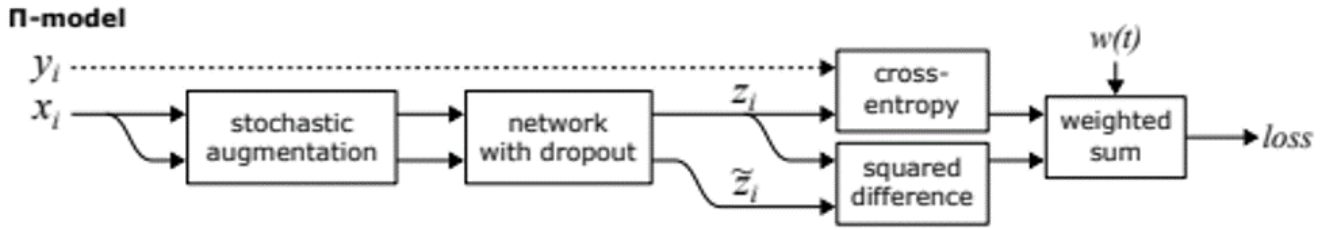


Fig 2: Π-MODEL Diagram Illustration

Temporal Ensembling Model

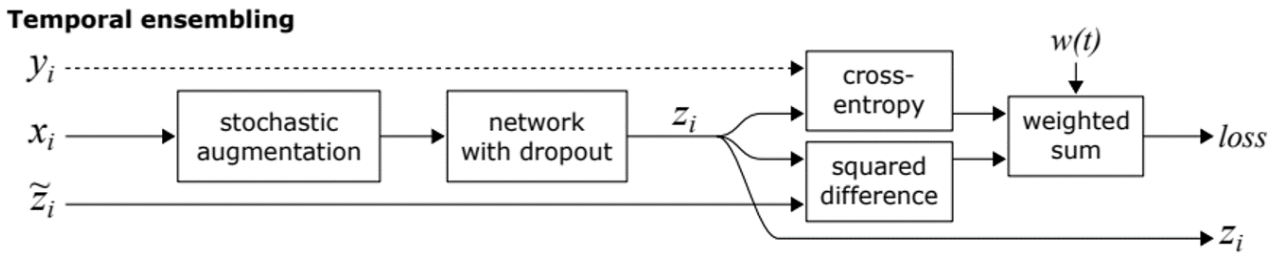


Fig 3: Temporal Ensembling Model Diagram Illustration

- *Advances the model by taking into account the network predictions over multiple previous training epochs*
- *Differs from the ∏ model in the following ways*
  - Dropout and Augmentations are evaluated just once for every input per epoch
  - and the target vectors for the unsupervised loss component are based on prior network evaluations
  - Faster and less noisy than the ∏ model

Temporal Ensembling Model

- Evaluate network outputs for augmented inputs
  - Loss Function is the sum of two components:

    Component 1. Standard cross entropy loss for labelled

    Component 2. Weighted Mean squared difference of the two prediction vectors (for all inputs)

  - Update Network parameters

  - Construct the target vectors which is then used in training the network on the same inputs for the unsupervised loss component.

  - Accumulate ensemble predictions.

Data Preprocessing

Data Preprocessing Stages

Dataset Information

- We use the sentiment140 dataset. It contains tweets extracted using the twitter API.

Data Set Characteristics:	Tweet messages	Number of Instances:	1,612,839	Area:	Social Media
Attribute Characteristics:	Categorical	Number of Attributes:	varies	Labeled data	12,839
Associated Tasks:	Classification	Missing Values?	nil	Unlabeled data	1,600,000

- **Source:** https://www.kaggle.com/kazanova/sentiment140

Data Preprocessing Stages

Libraries used

- Some of the libraries used for the data preprocessing includes NumPy, Pandas, matplotlib, Gensim, csv, os, tensorflow, etc

Extracting relevant information(tweets) from data

❖ *Pandas was used to extract tweets(unlabeled) from our dataset*

❖ *Shell script was used to remove irrelevant information from tweets (links and usernames)*

0	1467810369	Mon Apr 06 22:19:45 PDT 200	NO_QUERY	_TheSpecialOne_
1	1467810672	Mon Apr 06 22:19:49 PDT 200	NO_QUERY	scotthamilton
2	1467810917	Mon Apr 06 22:19:53 PDT 200	NO_QUERY	mattycus
3	1467811184	Mon Apr 06 22:19:57 PDT 200	NO_QUERY	ElleCTF
4	1467811193	Mon Apr 06 22:19:57 PDT 200	NO_QUERY	Karoli
5	1467811372	Mon Apr 06 22:20:00 PDT 200	NO_QUERY	joy_wolf
6	1467811592	Mon Apr 06 22:20:03 PDT 200	NO_QUERY	mybirch
7	1467811594	Mon Apr 06 22:20:03 PDT 200	NO_QUERY	coZZ
8	1467811795	Mon Apr 06 22:20:05 PDT 200	NO_QUERY	2Hood4Hollywood
9	1467812025	Mon Apr 06 22:20:09 PDT 200	NO_QUERY	mimismo
10	1467812416	Mon Apr 06 22:20:16 PDT 200	NO_QUERY	erinx3leannexo
11	1467812579	Mon Apr 06 22:20:17 PDT 200	NO_QUERY	pardonlauren
12	1467812723	Mon Apr 06 22:20:19 PDT 200	NO_QUERY	TLeC
13	1467812771	Mon Apr 06 22:20:19 PDT 200	NO_QUERY	robrobbierobert
14	1467812784	Mon Apr 06 22:20:20 PDT 200	NO_QUERY	bayofwolves
15	1467812799	Mon Apr 06 22:20:20 PDT 200	NO_QUERY	HairByJess
16	1467812964	Mon Apr 06 22:20:22 PDT 200	NO_QUERY	lovesongwriter
17	1467813137	Mon Apr 06 22:20:25 PDT 200	NO_QUERY	armotley
18	1467813579	Mon Apr 06 22:20:31 PDT 200	NO_QUERY	starkissed

@switchfoot http://twitpic.com/2y1zl - Awww, that's a bummer. You shoulda got David Carr of Third Day to do it. ;D
is upset that he can't update his Facebook by texting it... and might cry as a result School today also. Blah!
@Kenichan I dived many times for the ball. Managed to save 50% The rest go out of bounds
my whole body feels itchy and like its on fire
@nationwideclass no, it's not behaving at all. i'm mad. why am i here? because I can't see you all over there.
@Kwesidei not the whole crew
Need a hug
@LOLTrish hey long time no see! Yes.. Rains a bit ,only a bit LOL , I'm fine thanks , how's you ?
@Tatiana_K nope they didn't have it
@twittera que me muera ?
spring break in plain city... it's snowing
I just re-pierced my ears
@caregiving I couldn't bear to watch it. And I thought the UA loss was embarrassing . . . . .
@octolinz16 It it counts, idk why I did either. you never talk to me anymore
@smarrison i would've been the first, but i didn't have a gun. not really though, zac snyder's just a doucheclown.
@iamjazzyfizzle I wish I got to watch it with you!! I miss you and @iamlilnicki how was the premiere?!
Hollis' death scene will hurt me severely to watch on film wry is directors cut not out now?
about to file taxes
@LettyA ahh ive always wanted to see rent love the soundtrack!!

- Awww, that's a bummer. You shoulda got David Carr of Third Day to do it. ;D
is upset that he can't update his Facebook by texting it... and might cry as a result School today also. Blah!
I dived many times for the ball. Managed to save 50% The rest go out of bounds
my whole body feels itchy and like its on fire
no, it's not behaving at all. i'm mad. why am i here? because I can't see you all over there.
not the whole crew
Need a hug
hey long time no see! Yes.. Rains a bit ,only a bit LOL , I'm fine thanks , how's you ?
nope they didn't have it
que me muera ?
spring break in plain city... it's snowing
I just re-pierced my ears
I couldn't bear to watch it. And I thought the UA loss was embarrassing . . . . .

Data Preprocessing Stages

Covert categorical labels to numeric

- Used NumPy to convert the target features from categorical to numeric values

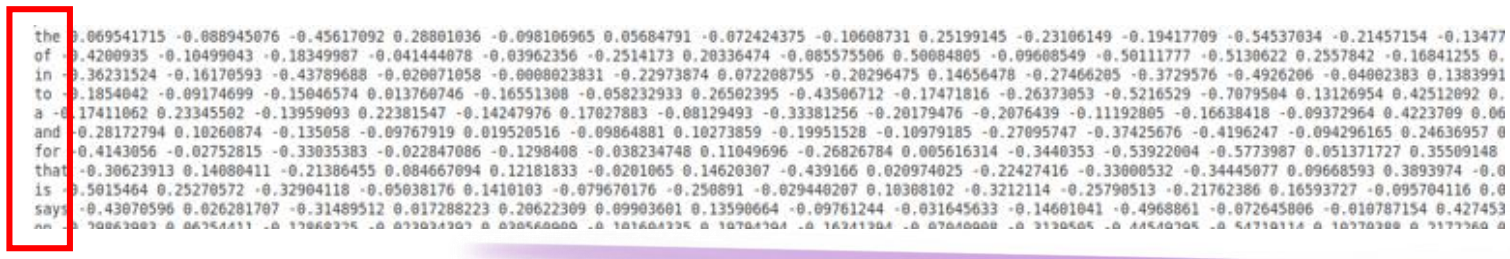
Statement	label
Building a wall on the U.S.-Mexico border will take literally years.	1
Wisconsin is on pace to double the number of layoffs this year.	0
Says John McCain has done nothing to help the vets.	0
Suzanne Bonamici supports a plan that will cut choice for Medicare Advantage seniors.	2
When asked by a reporter whether hes at the center of a criminal scheme to violate campaign laws, Gov. Scott Walker nodded yes.	4
Over the past five years the federal government has paid out $601 million in retirement and disability benefits to deceased former federal employees.	1
Says that Tennessee law requires that schools receive half of proceeds -- $31 million per year -- from a half-cent increase in the Shelby County sales tax.	1
Says Vice President Joe Biden "admits that the American people are being scammed" with the economic stimulus package.	3
Donald Trump is against marriage equality. He wants to go back.	1
We know that more than half of Hillary Clintons meetings while she was secretary of state were given to major contributors to the Clinton Foundation.	3
We know there are more Democrats in Georgia than Republicans. We know that for a fact.	3
PolitiFact Texas says Congressman Edwards attacks on Bill Flores are false.	3
Denali is the Kenyan word for black power.	4
Says 57 percent of federal spending goes to the military and just 1 percent goes to food and agriculture, including food stamps.	0
On residency requirements for public workers	2
Says the unemployment rate for college graduates is 4.4 percent and over 10 percent for noncollege-educated.	1
Unfortunately we have documented instances where people defecated in the (Statehouse) building.	4
A recent Gallup poll found that 72 percent of Americans and 56 percent of Democrats say the biggest threat to our nations security is big government.	2
"Each year, 18,000 people die in America because they don't have health care."	1

❖ {"false": 0, "true": 1, "half-true": 2, "barely-true": 3, "pants-fire": 4, "mostly-true": 5}

Convert corpus (tweets) to vectors

❖ Using the 'word2vec' from the Gensim model, we convert the corpus(tweets) to vectors using Gensim

❖ Specific dimensions (100) and window size (5)



Data Preprocessing Stages

Mapped the vector to matrix

- We represent each tweet by a matrix whose dimensions are "words" by "vectors" using NumPy

```
the 0.069541715 -0.088945076 -0.45617092 0.28801036 -0.098106965 0.05684791 -0.072424375 -0.10608731 0.25199145 -0.23106149 -0.19417709 -0.54537034 -0.21457154 -0.13477
of -0.4200935 -0.10499043 -0.18349987 -0.041444078 -0.03962356 -0.2514173 0.20336474 -0.085575506 0.50084805 -0.09608549 -0.50111777 -0.5130622 0.2557842 -0.16841255 0.
in -0.36231524 -0.16170593 -0.43789688 -0.020071058 -0.0008023831 -0.22973874 0.072208755 -0.20296475 0.14656478 -0.27466205 -0.3729576 -0.4926206 -0.04002383 0.1383991
to -0.1854042 -0.09174699 -0.15046574 0.013760746 -0.16551308 -0.058232933 0.26502395 -0.43506712 -0.17471816 -0.26373053 -0.5216529 -0.7079504 0.13126954 0.42512092 0.
a -0.17411062 0.23345502 -0.13959093 0.22381547 -0.14247976 0.17027883 -0.08129493 -0.33381256 -0.20179476 -0.2076439 -0.11192805 -0.16638418 -0.09372964 0.4223709 0.06
and -0.28172794 0.10260874 -0.135058 -0.09767919 0.019520516 -0.09864881 0.10273859 -0.19951528 -0.10979185 -0.27095747 -0.37425676 -0.4196247 -0.094296165 0.24636957 0
for -0.4143056 -0.02752815 -0.33035383 -0.022847086 -0.1298408 -0.038234748 0.11049696 -0.26826784 0.005616314 -0.3440353 -0.53922004 -0.5773987 0.051371727 0.35509148
that -0.30623913 0.14080411 -0.21386455 0.084667094 0.12181833 -0.0201065 0.14620307 -0.439166 0.020974025 -0.22427416 -0.33000532 -0.34445077 0.09668593 0.3893974 -0.0
is -0.5015464 0.25270572 -0.32904118 -0.05038176 0.1410103 -0.079670176 -0.250891 -0.029440207 0.10308102 -0.3212114 -0.25790513 -0.21762386 0.16593727 -0.095704116 0.0
says -0.43070596 0.026281707 -0.31489512 0.017288223 0.20622309 0.09903601 0.13590664 -0.09761244 -0.031645633 -0.14601041 -0.4968861 -0.072645806 -0.010787154 0.427453
```

Covert the data to 'tfrecord' format

❖ We finally convert the data to tensor flow data format for the neural network training

Experimental Results

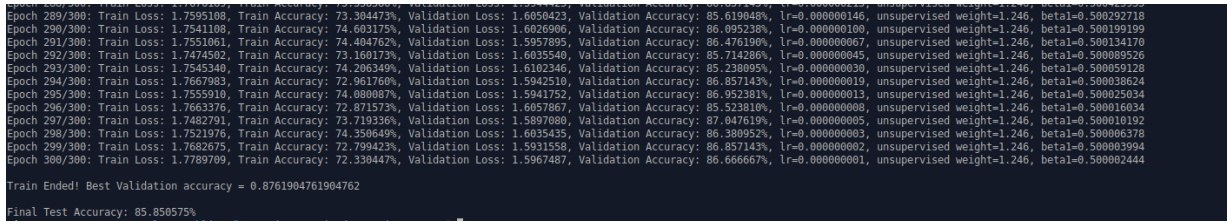
SVHN Dataset

- *Still at the final stages of preprocessing of the twitter dataset*
- *For purposes of the experiment results, we use the SVHN dataset to train and test the model*
- *The street view house numbers (SVHN) dataset consists of 32 × 32 pixel RGB images of real-world house numbers*
- *Task is to classify the centermost digit.*

Dataset	Domain	Dimensionality	Training set	Test set
SVHN	Vision	3072(32 x 32)	73k	26k

Results after 300 Epochs

- **Training Accuracy – 72.33% Training Duration: 10 hrs**
- **Validation Accuracy – 0.876**
- **Final Test Accuracy – 85.85%**



Method	Epoch	Train. Acc.	Valid Acc.	Test. Acc.
Temporal Ensembling	300	72.33%	87.6%	85.85%

Future work

- Further work on this project would be to test the Temporal Ensembling model on the preprocessed twitter dataset, evaluate the performance and possible carry out research to improve the model performance in fake news detection.
- Research possible methods or ways to improve the time taken in training the model.

Conclusion

A major part of this project was dedicated to data preprocessing as the data had to be converted to a format that our model can communicate with. The next step will be to test our model using our new dataset and seek further ways to improve the model

THANK YOU

Q & A