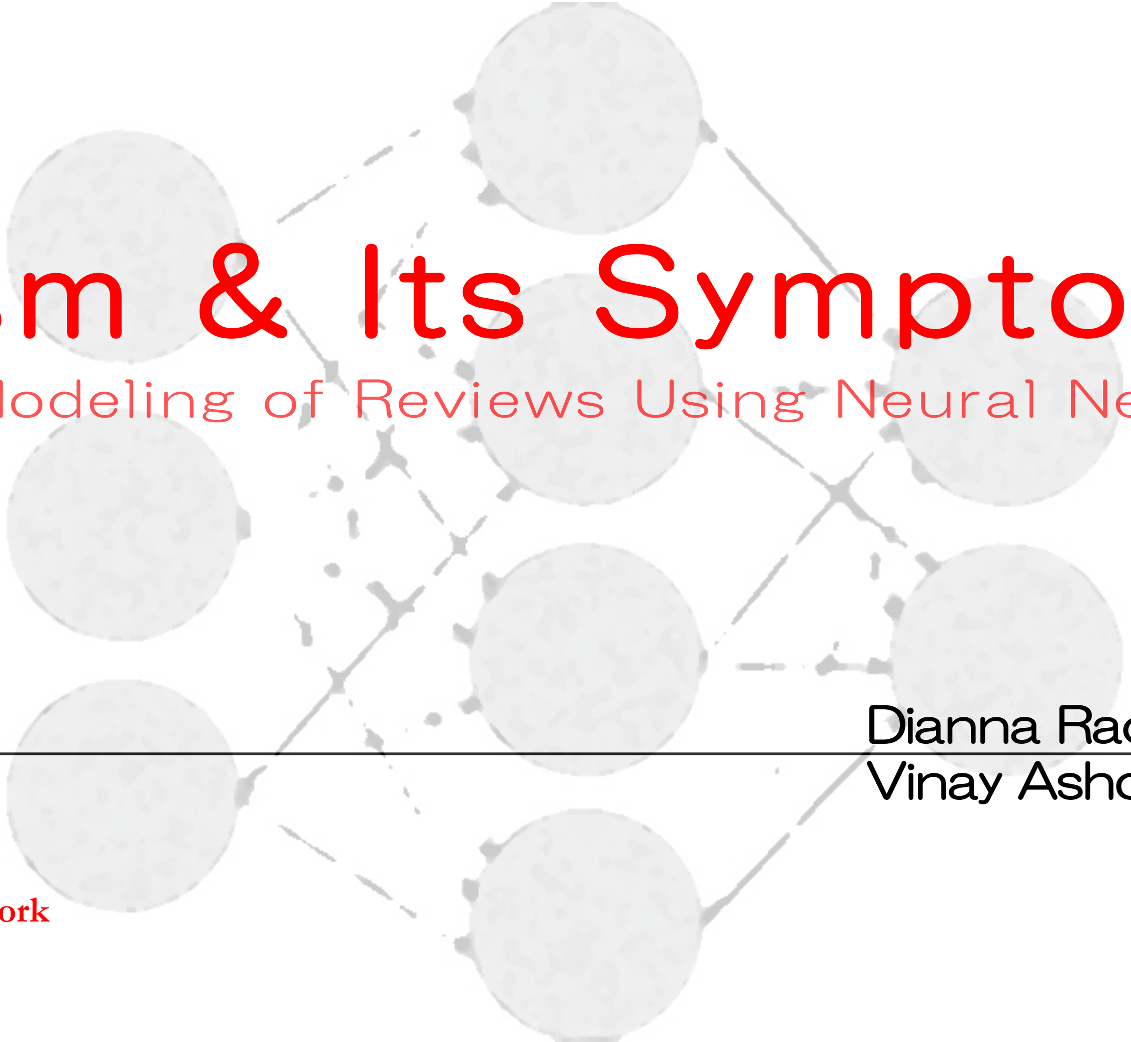
Sarcasm & Its Symptoms

Rating Based Modeling of Reviews Using Neural Networks

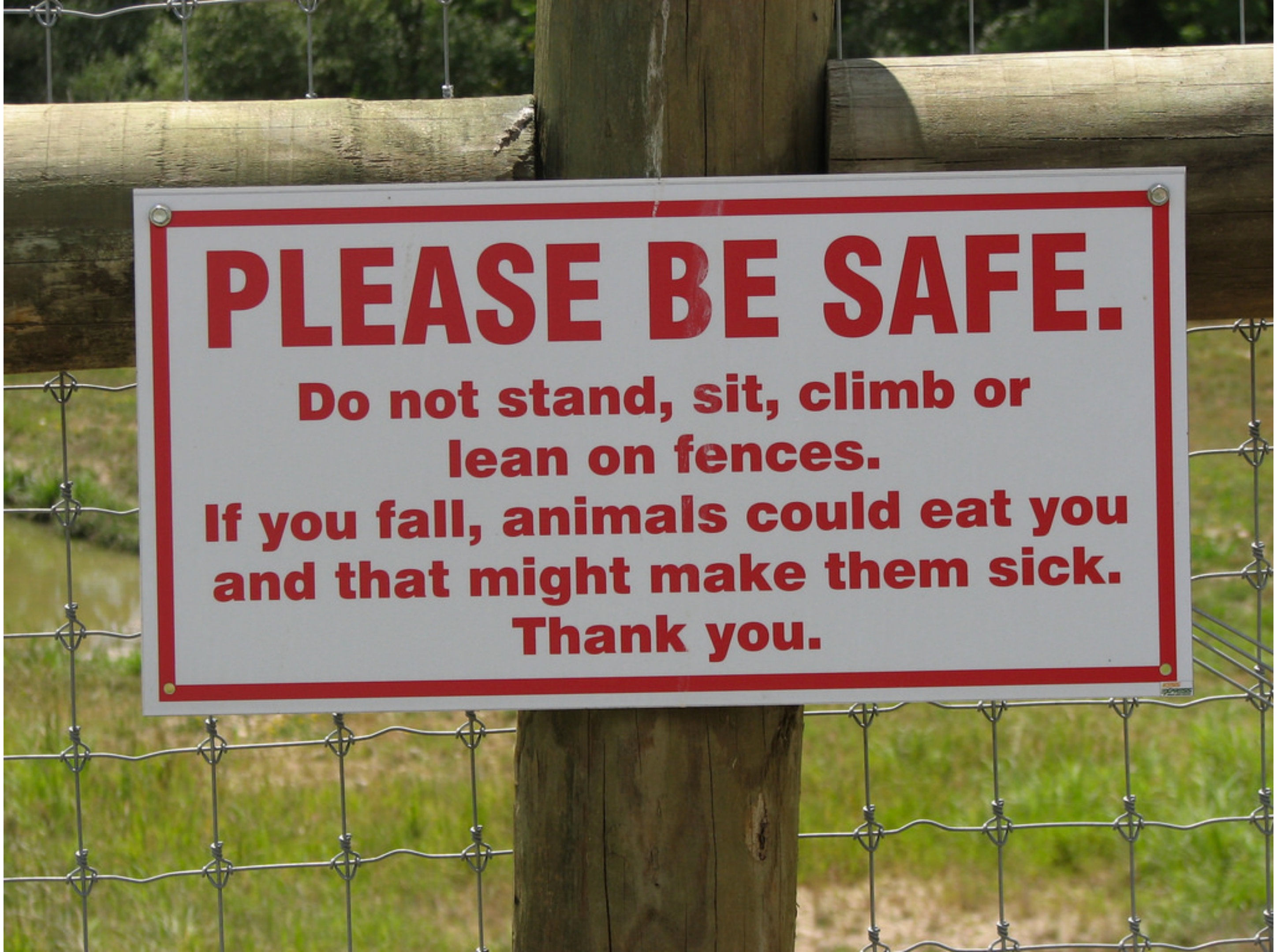


Dianna Radpour
Vinay Ashokkumar

University at Buffalo
The State University of New York

Aim: Sarcasm detection using deep learning

○ a method of classifying sarcastic text on a purely non-contextual basis, i.e. inferences and through only impromptu presence of the text and features



PLEASE BE SAFE.

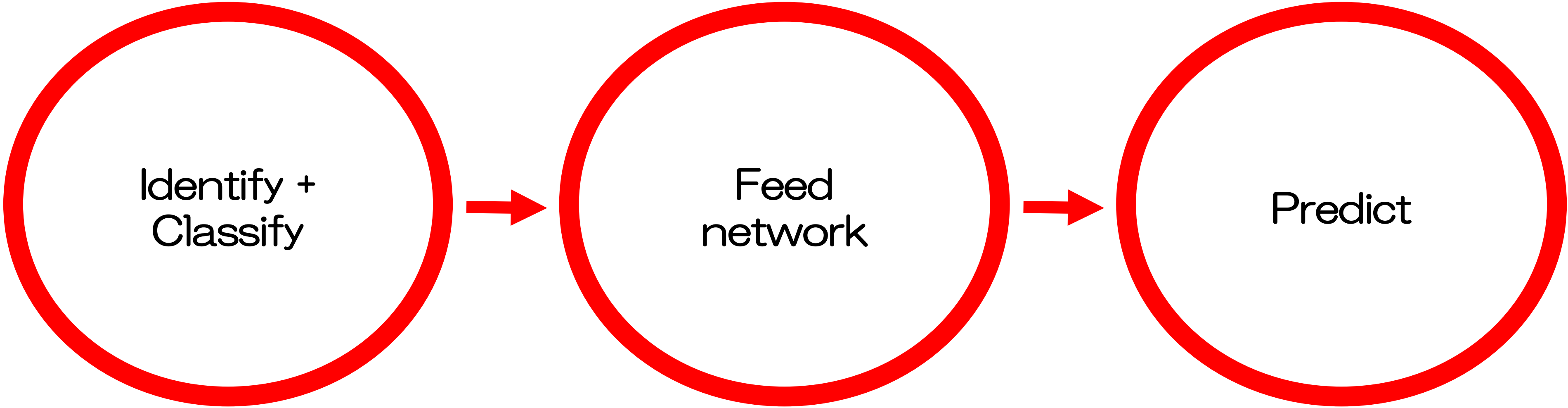
Do not stand, sit, climb or
lean on fences.
If you fall, animals could eat you
and that might make them sick.
Thank you.

Relevant Queries

- Are users more likely to be more or less sarcastic during the holiday season?

- Does higher usage of exclamation points indicate a sarcastic review?

- Do businesses with lower ratings tend to have more sarcastic reviews?

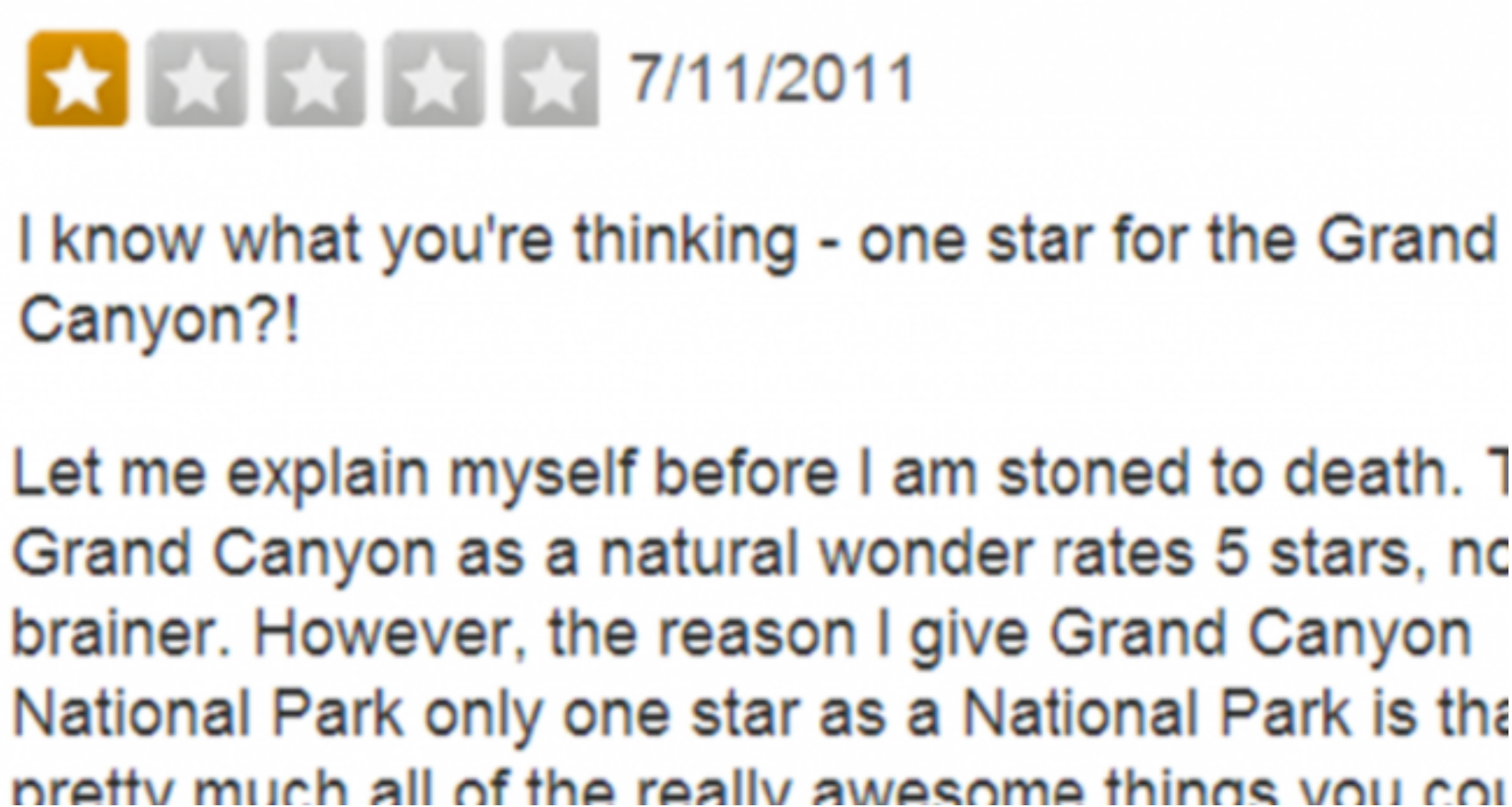


Identify + Classify

Feed network

Predict

Identification, i.e. sarcasm detection, is one of the most challenging aspects of sentiment analysis.



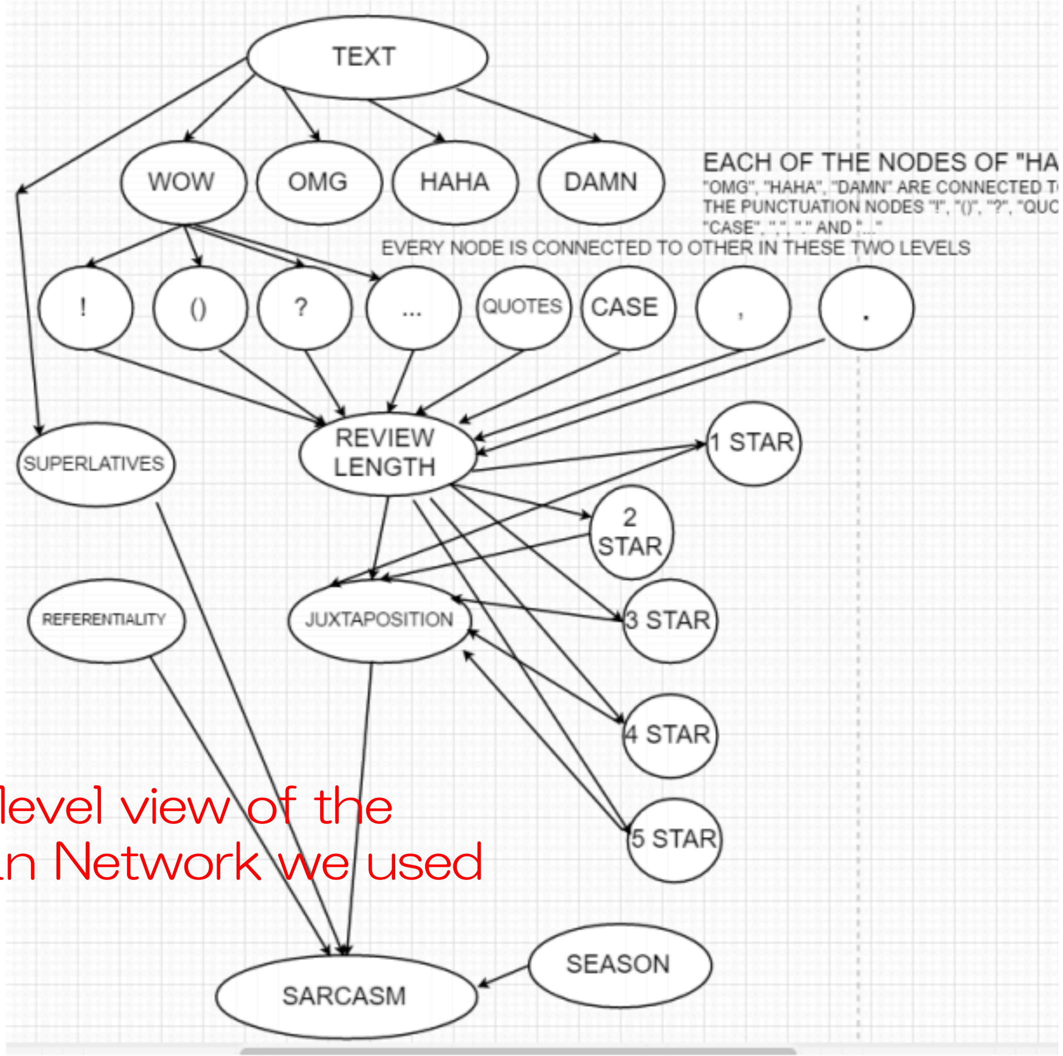
★☆☆☆☆ 7/11/2011

I know what you're thinking - one star for the Grand Canyon?!

Let me explain myself before I am stoned to death. T
Grand Canyon as a natural wonder rates 5 stars, nc
brainer. However, the reason I give Grand Canyon
National Park only one star as a National Park is tha
pretty much all of the really awesome things you cou

cutting
Yacob
question
San
in
Cantonese
understand
according
tool
written
gibe
situation
dementia
brain
rage
statement
certain
posted
later
sarcasm.[22]A
mark
Huang
Dutch
people
expresser
intention
theoretical
19th
unless
Galinsky
forms
Tsigie
generally
Sarcasm
use
Main
years
seriously
find
fine
gave
First
late
One
cue
since
eyes
School
several
linguistic
now
tends
indication
raising
like
analytics
bite
sarcastic
necessarily
aid
lead
avoid
uses
terms
travel
slaq in
extent
real
hand
irony
o
pitch
shows
treat
mumble
offensive
soul
A.D
groups
tests
thus
conversation
last
well
thereafter
found
e.g
will
cry
wits
good
warning
end
many
see
novel
mere
Brant
ironic
long
vary
directly
also
Gino
lip
bids
within
place
Harvard
face
may
form
although
activate
free
media
Tom
model
one
different
entry
two
bitter
can
ways
used
tone
point
careful
devil
ideas
word
meant
must
lower
sentence
turned
modest
executives
often
expression
suggest
slaq
construction
mass
speakers
intercultural
comment
play
understanding
Well
meaning
test
tear
the University
first
said
unicode
content
show
feel
proposal
usage
sophisticated
difficulties
widely
interpretation
reduced
language
things
translating
taken
sometimes
refuge
include
expressions
privacy
meetings
without
lowered
communications
overt
expressed
research
rising
reason
recipient
Lazarus
across
greater
lying
August
pain
spoken
finding
denote
English
intentions
Business
slowly
marks
study
insulted
recent
North
one's
early
much
present
politeness
emotional
indicated
damage
requires
upon
June
California
especially
looks
Also
government
Secret
age
exclamation
business
remark
neutral
varying
vocal
manner
signal
Daniel
told
company
punctuation
individual
context
cultures

Sarcastic Feature Selection


TEXT
WOW
OMG
HAHA
DAMN
EACH OF THE NODES OF "HA
"OMG", "HAHA", "DAMN" ARE CONNECTED T
THE PUNCTUATION NODES "!", "()", "?", "QUO
"CASE", ",", "." AND "..."
EVERY NODE IS CONNECTED TO OTHER IN THESE TWO LEVELS
!
()
?
...
QUOTES
CASE
,
.
SUPERLATIVES
REVIEW LENGTH
1 STAR
2 STAR
REFERENTIALITY
JUXTAPOSITION
3 STAR
4 STAR
5 STAR
SEASON
SARCASM

a high-level view of the Bayesian Network we used

Parameters of the Model

Feature	Types
Keyword(s)	“wow” “OMG” “damn”
Punctuation	“!” “()” “…” “ “” ”
Referentiality	“I” “me” “myself” “we” “us”
Season	“winter” “fall” “summer” “spring”
Superlatives	“worst” “best” “word ending in -est”

Method

- harness the presence of contextual incongruity in user reviews

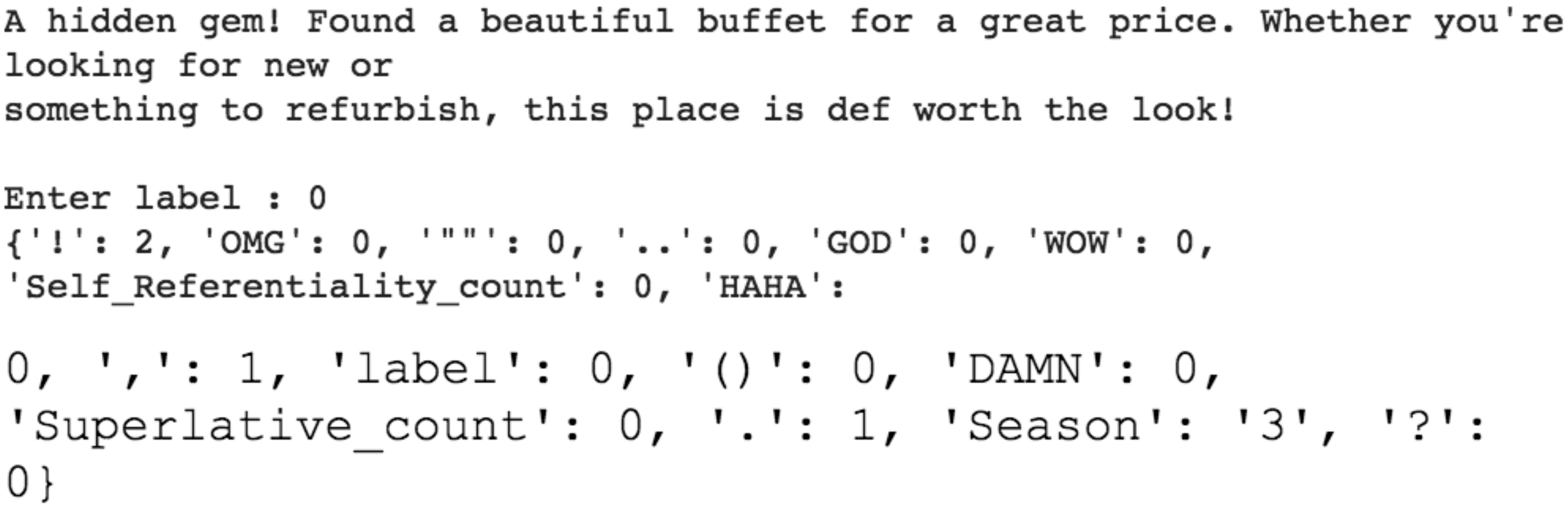
  i.e. sentiment juxtaposition (5-star in rating with a negative sentiment, & vice versa)

inconsistency in sentiment will be indicative of some level of ironic text

Data

preprocessing step is initialized with creating a dictionary that reads through each of the (~4mil) reviews and calculates the # of times for each instance of a categorized feature appearing in that review text

An example from a 4 star review

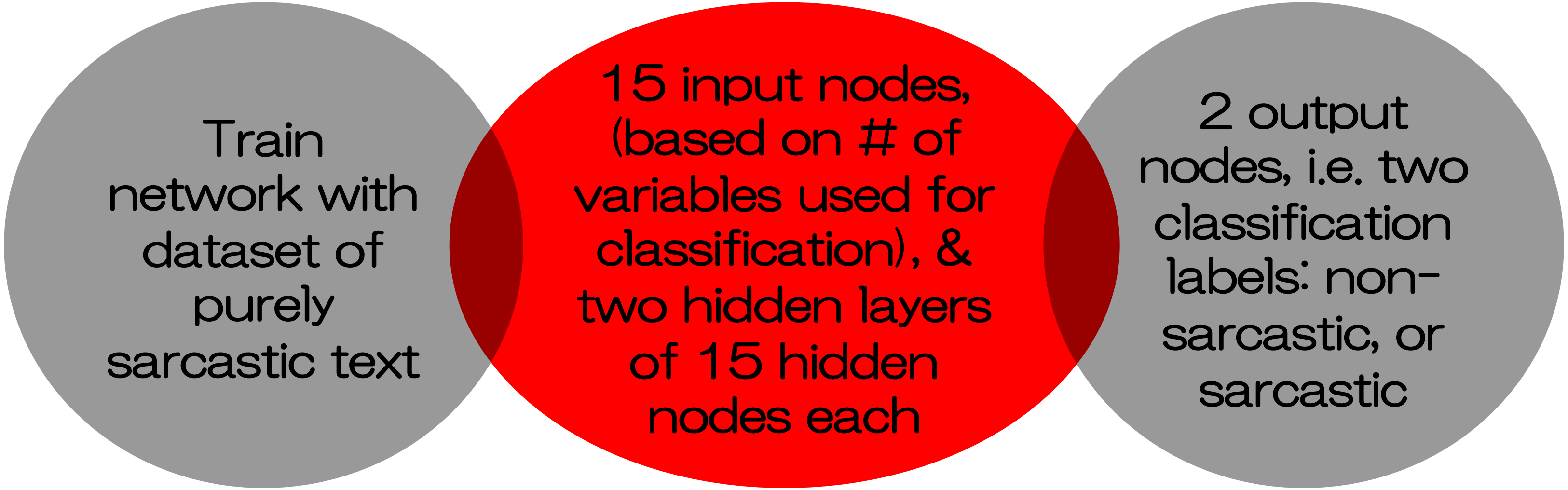


```
A hidden gem! Found a beautiful buffet for a great price. Whether you're
looking for new or
something to refurbish, this place is def worth the look!

Enter label : 0
{'!': 2, 'OMG': 0, '""': 0, '..': 0, 'GOD': 0, 'WOW': 0,
'Self_Referentiality_count': 0, 'HAHA':
0, ',': 1, 'label': 0, '()': 0, 'DAMN': 0,
'Superlative_count': 0, '.': 1, 'Season': '3', '?':
0}
```

THIS IS
DEEPLEARNING

A simple network



Train network with dataset of purely sarcastic text

15 input nodes, (based on # of variables used for classification), & two hidden layers of 15 hidden nodes each

2 output nodes, i.e. two classification labels: non-sarcastic, or sarcastic

1-star and 3-star rated reviews demonstrated a high level of accuracy, 95% and 96% respectively

Notes on network parameters

built-in optimizer (Adam) used to calculate the cross entropy between the predicted & true label

Activation functions used between pairs of levels (Relu for the path from the 1st to the 2nd hidden layer)

applications of deep learning algorithms

Why do we want to know if users are being sarcastic?

- less genuinely indicative of customer attitude
- would definitely get some more valuable feedback on the company at large
- will help to improve the customer experience at large.

BEWARE
OF
SARCASM

Future Directions

- use more manpower to rate the sarcastic nature of the reviews (rather than just us 2)
- Implementation of different machine learning models e.g. Restricted Boltzmann Machines and Hidden Markov Models to generate a time series model for for feature analysis of sarcasm in *real time*

Thanks!

Questions?

diannara@buffalo.edu