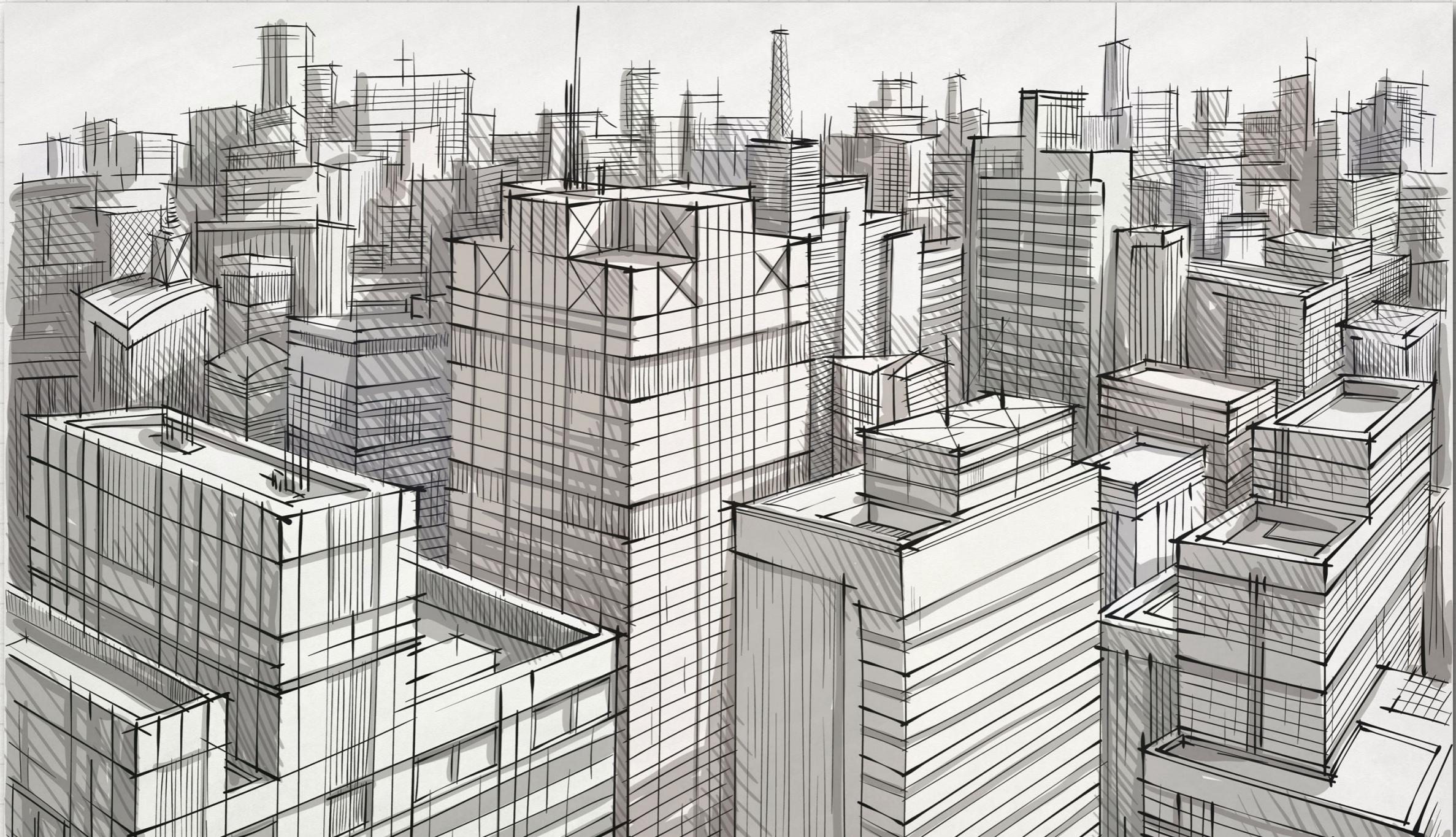


EASE - EASY AUTOMATED SCORING OF ESSAY

ANUJ CHAUHAN



AGENDA

- Motivation
- Problem Statement
- Data Description
- Challenge
- Approach
- Results
- Conclusions

MOTIVATION

- Automation
- Hypothesis:
How well can we teach a model to do a subjective human task.
- Apply skills and techniques from this semester.

- Motivation
- **Problem Statement**
- Data Description
- Challenge
- Approach
- Results
- Conclusions

PROBLEM STATEMENT

- Grade Essays
- Varying nature of essays
- Scored by human raters, then resolved by aggregation.

- Motivation
- Problem Statement
- **Data Description**
- Challenge
- Approach
- Results
- Conclusions

DATA DESCRIPTION

- Obtained from Kaggle.
- 8 essay sets.
- Different rater scores, one resolved score.
- Mostly clean, fortunately.
- Selected essay set 1,3,5,6.

essay_set	count
1	1783
2	1800
3	1726
4	1772
5	1805
6	1800
7	1569
8	723

DATA DESCRIPTION

- First few rows look like:

essay_id	essay_set	essay	rater1_domain1	rater2_domain1	domain1_score
0	1	1 Dear local newspaper, I think effects computer...	4.0	4.0	8.0
1	2	1 Dear @CAPS1 @CAPS2, I believe that using compu...	5.0	4.0	9.0
2	3	1 Dear, @CAPS1 @CAPS2 @CAPS3 More and more peopl...	4.0	3.0	7.0
3	4	1 Dear Local Newspaper, @CAPS1 I have found that...	5.0	5.0	10.0
4	5	1 Dear @LOCATION1, I know having computers has a...	4.0	4.0	8.0
5	6	1 Dear @LOCATION1, I think that computers have a...	4.0	4.0	8.0
6	7	1 Did you know that more and more people these d...	5.0	5.0	10.0

- A sample essay:

"Dear @CAPS1 @CAPS2, I believe that using computers will benefit us in many ways like talking and becoming friends with others through websites like facebook and mysace. Using computers can help us find coordinates, locations, and able ourselves to millions of information. Also computers will benefit us by helping with jobs as in planning a house plan and typing a @NUM1 page report for one of our jobs in less than writing it. Now lets go into the wonder world of technology. Using a computer will help us in life by talking or making friends on line. Many people have myspace, facebooks, aim, these all benefit us by having conversations with one another. Many people believe computers are bad but how can you make friends if you can never talk to them? I am very fortunate for having a computer that can help with not only school work but my social life and how I make friends. Computers help us with finding our locations, coordinates and millions of information online. If we didn't go on the internet a lot we wouldn't know how to go onto websites that @MONTH1 help us with locations and coordinates like @LOCATION1. Would you rather use a computer or be in @LOCATION3. When your supposed to be vacationing in @LOCATION2. Million of information is found on the internet. You can ask almost every question and a computer will have it. Would you rather easily draw up a house plan on the computers or take @NUM1 hours doing one by hand with ugly eraser marks all over it, you are guaranteed that to find a job with a drawing like that. Also when applying for a job many workers must write very long papers like a @NUM3 word essay on why this job fits you the most, and many people I know don't like writing @NUM3 words non-stop for hours when it could take them I have an a computer. That is why computers we needed a lot now adays. I hope this essay has impacted your decision on computers because they are great machines to work with. The other day I showed my mom how to use a computer and she said it was the greatest invention sense sliced bread! Now go out and buy a computer to help you chat online with friends, find locations and millions of information on one click of the button and help your self with getting a job with neat, prepared, printed work that your boss will love."

- Motivation
- Problem Statement
- Data Description
- Challenge
- Approach
- Results
- Conclusions

CHALLENGE

"Dear @CAPS1 @CAPS2, I believe that using computers will benefit us in many ways like talking and becoming friends with others through websites like facebook and myspace. Using computers can help us find coordinates, locations, and abilities ourselves to millions of information. Also computers will benefit us by helping with jobs as in planning a house plan and typing a @NUM1 page report for one of our jobs in less than writing it. Now lets go into the wonder world of technology. Using a computer will help us in life by talking or making friends on line. Many people have myspace, facebook, aim, these all benefit us by having conversations with one another. Many people believe computers are bad but how can you make friends if you can never talk to them? I am very fortunate for having a computer that can help with not only school work but my social life and how I make friends. Computers help us with finding our locations, coordinates and millions of information online. If we didn't go on the internet a lot we wouldn't know how to go onto websites that @MONTH1 help us with locations and coordinates like @LOCATION1. Would you rather use a computer or be in @LOCATION3. When your supposed to be vacationing in @LOCATION2. Millions of information is found on the internet. You can ask almost every question and a computer will have it. Would you rather easily draw up a house plan on the computers or take @NUM1 hours doing one by hand with ugly eraser marks all over it, you are guaranteed that to find a job with a drawing like that. Also when applying for a job many workers must write very long papers like a @NUM3 word essay on why this job fits you the most, and many people I know don't like writing @NUM3 words non-stop for hours when it could take them I have an a computer. That is why computers we needed a lot now adays. I hope this essay has impacted your decision on computers because they are great machines to work with. The other day I showed my mom how to use a computer and she said it was the greatest invention since sliced bread! Now go out and buy a computer to help you chat online with friends, find locations and millions of information on one click of the button and help your self with getting a job with neat, prepared, printed work that your boss will love."

8.0



'Dear Local Newspaper, @CAPS1 I have found that many experts say that computers do not benefit our society. In some cases this is true but in most cases studies show that computers can help people. While nothing beats talking in person computers can get really close such examples are webcams or e-mail. @PERCENT1 of students who get good grades refer to reliable websites for research or to help find good books. Also online catalogues or advertisements help the economy of stores worldwide. @CAPS2 people were not allowed to use computers most of the modern would not exist. @PERSON1 said that the best form of modern communication is the computer because of the ability to write, talk, or write back for much cheaper! Almost every single event I go to is planned on a computer by communication such as e-mail "@CAPS2 a student ever needs homework because I am sick or needs help studying for a test then contact their teacher through the best form of communication for them always e-mail. Even the post office uses computers to get letters and boxes to people. The president of the post office, @PERSON2 said "@CAPS3 would be impossible to get mail to our customers @CAPS2 @CAPS3 were not for computers telling us where a zip code is or how heavy a box is." @CAPS4 that tell people what is happening around the world would not exist @CAPS2 @CAPS3 were not for the modern communication abilities that computer provides. Because information can be communicated so quickly, so can research. When the country of @LOCATION2 took a poll @PERCENT2 of people used computer for any type of research, of those @PERCENT3 were students currently in school and @PERCENT4 of them have good grades. When the same survey was taken in the @LOCATION1 @PERCENT5 of people used computers for reasons and @PERCENT2 were students who had good grades @CAPS2 @CAPS3 were not possible for me to access documents in the @CAPS5 @CAPS6 online I probably would not have gotten an A+ on my @CAPS7 assignment! Could you imagine @CAPS2 suddenly your Newspaper reporters couldn't use the internet to work on their reports? The articles would probably be @NUM1 after the events occur. Most business, including the Newspaper, use the internet to advertise, shop, or read. The association of @ORGANIZATION1 reported that in @PERCENT1 of @ORGANIZATION1 used a website and of them @PERCENT5 were in good positions. The president of @CAPS8 @NUM2 imports said that they use an online catalogue because @CAPS3 is cheaper, but they can also promote that @CAPS3 is to save trees, or for the customer's convenience. Small @ORGANIZATION1 can make websites to promote themselves and explain their story to potential customers. @PERSON3, the owner of @ORGANIZATION2's said that the internet saved her restaurant. @CAPS2 @CAPS3 were not for the internet @NUM3 more people would be jobless in @LOCATION3. In conclusion computer help everyday people and without them most conveniences would not exist. They help communicate around the world. Computers help people research subjects for school reports, and they make the current economy get better everyday. In moderation computers are the most useful tool out there.'

9.0



"Dear, @CAPS1 @CAPS2 @CAPS3 More and more people use computers, but not everyone agrees that this benefits society. Those who support advances in technology believe that computers have a positive effect on people. Others have different ideas. A great amount in the world today are using computers, some for work and some for the fun of it. Computers is one of man's greatest accomplishments. Computers are helpful in so many ways, @CAPS4, news, and live streams. Don't get me wrong way too much people spend time on the computer and they should be out interacting with others but who are we to tell them what to do. When I grow up I want to be a author or a journalist and I know for a fact that both of those jobs involve lots of time on the computer, one @MONTH1 spend more time than the other but you know exactly what @CAPS5 getting at. So what if some experts think people are spending too much time on the computer and not exercising, enjoying nature and interacting with family and friends. For all the experts know that it's how most people make a living and we don't know why people choose to use the computer for a great amount of time and to be honest it's none of my concern and it shouldn't be the so called experts concern. People interact a thousand times a day on the computers. Computers keep lots of kids off the streets instead of being out and causing trouble. Computers help the @ORGANIZATION1 locate most wanted criminals. As you can see computers are more useful to society than you think, computers benefit society."

7.0

- Motivation
- Problem Statement
- Data Description
- Challenge
- Approach
- Results
- Conclusions

APPROACH

FEATURE GENERATION

- Generate numerical features.

META FEATURES

1. Count of words.
2. Count of sentences.
3. Count of distinct words.
4. Count of syllables.

"Dear @CAPS1 @CAPS2, I believe that using computers will benefit us in many ways like talking and becoming friends with others through websites like facebook and myspace. Using computers can help us find coordinates, locations, and able ourselves to millions of information. Also computers will benefit us by helping with jobs as in planning a house plan and typing a @NUM1 page report for one of our jobs in less than writing it. Now lets go into the wonder world of technology. Using a computer will help us in life by talking or making friends on line. Many people have myspace, facebooks, aim, these all benefit us by having conversations with one another. Many people believe computers are bad but how can you make friends if you can never talk to them? I am very fortunate for having a computer that can help with not only school work but my social life and how I make friends. Computers help us with finding our locations, coordinates and millions of information online. If we didn't go on the internet a lot we wouldn't know how to go onto websites that @MONTH1 help us with locations and coordinates like @LOCATION1. Would you rather use a computer or be in @LOCATION3. When your supposed to be vacationing in @LOCATION2. Million of information is found on the internet. You can ask almost every question and a computer will have it. Would you rather easily draw up a house plan on the computers or take @NUM1 hours doing one by hand with ugly eraser marks all over it, you are guaranteed that to find a job with a drawing like that. Also when applying for a job many workers must write very long papers like a @NUM3 word essay on why this job fits you the most, and many people I know don't like writing @NUM3 words non-stop for hours when it could take them I have an a computer. That is why computers we needed a lot now adays. I hope this essay has impacted your decision on computers because they are great machines to work with. The other day I showed my mom how to use a computer and she said it was the greatest invention sense sliced bread! Now go out and buy a computer to help you chat online with friends, find locations and millions of information on one click of the button and help your self with getting a job with neat, prepared, printed work that your boss will love."

APPROACH

FEATURE GENERATION

- Generate numerical features.

"Dear @CAPS1 @CAPS2, I believe that using computers will benefit us in many ways like talking and becoming friends with others through websites like facebook and myspace. Using computers can help us find coordinates, locations, and allow ourselves to millions of information. Also computers will benefit us by helping with jobs as in planning a house plan and typing a @NUM1 page report for one of our jobs in less than writing it. Now lets go into the wonder world of technology. Using a computer will help us in life by talking or making friends online. Many people have myspace, facebooks, aim, these all benefit us by having conversations with one another. Many people believe computers are bad but how can you make friends if you can never talk to them? I am very fortunate for having a computer that can help with not only school work but my social life and how I make friends. Computers help us with finding our locations, coordinates and millions of information online. If we didn't go on the internet a lot we wouldn't know how to go onto websites that @MONTH1 help us with locations and coordinates like @LOCATION1. Would you rather use a computer or be in @LOCATION3. When your supposed to be vacationing in @LOCATION2. Millions of information is found on the internet. You can ask almost every question and a computer will have it. Would you rather easily draw up a house plan on the computers or take @NUM1 hours doing one by hand with ugly eraser marks all over it, you are guaranteed that to find a job with a drawing like that. Also when applying for a job many workers must write very long papers like a @NUM3 word essay on why this job fits you the most, and many people I know don't like writing @NUM3 words non-stop for hours when it could take them I have an a computer. That is why computers we needed a lot now adays. I hope this essay has impacted your decision on computers because they are great machines to work with. The other day I showed my mom how to use a computer and she said it was the greatest invention sense sliced bread! Now go out and buy a computer to help you chat online with friends, find locations and millions of information on one click of the button and help your self with getting a job with neat, prepared, printed work that your boss will love."

LEXICAL FEATURES

5. Count of Nouns.
6. Count of Proper nouns.
7. Count of Verbs.
8. Count of Adjectives.

APPROACH

FEATURE GENERATION

- Generate numerical features.

"Dear @CAPS1 @CAPS2, I believe that using computers will benefit us in many ways like talking and becoming friends with others through websites like facebook and myspace. Using computers can help us find coordinates, locations, and able ourselves to millions of information. Also computers will benefit us by helping with jobs as in planning a house plan and typing a @NUM1 page report for one of our jobs in less than writing it. Now lets go into the wonder world of technology. Using a computer will help us in life by talking or making friends online. Many people have myspace, facebooks, aim, these all benefit us by having conversations with one another. Many people believe computers are bad but how can you make friends if you can never talk to them? I am very fortunate for having a computer that can help with not only school work but my social life and how I make friends. Computers help us with finding our locations, coordinates and millions of information online. If we didn't go on the internet a lot we wouldn't know how to go onto websites that @MONTH1 help us with locations and coordinates like @LOCATION1. Would you rather use a computer or be in @LOCATION3. When your supposed to be vacationing in @LOCATION2. Millions of information is found on the internet. You can ask almost every question and a computer will have it. Would you rather easily draw up a house plan on the computers or take @NUM1 hours doing one by hand with ugly eraser marks all over it, you are guaranteed that to find a job with a drawing like that. Also when applying for a job many workers must write very long papers like a @NUM3 word essay on why this job fits you the most, and many people I know don't like writing @NUM3 words non-stop for hours when it could take them I have an a computer. That is why computers we needed a lot now adays. I hope this essay has impacted your decision on computers because they are great machines to work with. The other day I showed my mom how to use a computer and she said it was the greatest invention sense sliced bread! Now go out and buy a computer to help you chat online with friends, find locations and millions of information on one click of the button and help your self with getting a job with neat, prepared, printed work that your boss will love."

DERIVED READABILITY INDICES

9. Simple Readability Index
10. Flesch Reading Ease
11. Gunning Fog Index
12. Smog Index
13. Dale Chall Readability Index

APPROACH

FEATURE GENERATION

- Generate numerical features.

"Dear @CAPS1 @CAPS2, I believe that using computers will benefit us in many ways like talking and becoming friends with others through websites like facebook and myspace. Using computers can help us find coordinates, locations, and allow ourselves to millions of information. Also computers will benefit us by helping with jobs as in planning a house plan and typing a @NUM1 page report for one of our jobs in less than writing it. Now lets go into the wonder world of technology. Using a computer will help us in life by talking or making friends on line. Many people have myspace, facebooks, aim, these all benefit us by having conversations with one another. Many people believe computers are bad but how can you make friends if you can never talk to them? I am very fortunate for having a computer that can help with not only school work but my social life and how I make friends. Computers help us with finding our locations, coordinates and millions of information online. If we didn't go on the internet a lot we wouldn't know how to go onto websites that @MONTH1 help us with locations and coordinates like @LOCATION1. Would you rather use a computer or be in @LOCATION3. When your supposed to be vacationing in @LOCATION2. Millions of information is found on the internet. You can ask almost every question and a computer will have it. Would you rather easily draw up a house plan on the computers or take @NUM1 hours doing one by hand with ugly eraser marks all over it, you are guaranteed that to find a job with a drawing like that. Also when applying for a job many workers must write very long papers like a @NUM3 word essay on why this job fits you the most, and many people I know don't like writing @NUM3 words non-stop for hours when it could take them I have an a computer. That is why computers we needed a lot now adays. I hope this essay has impacted your decision on computers because they are great machines to work with. The other day I showed my mom how to use a computer and she said it was the greatest invention sense sliced bread! Now go out and buy a computer to help you chat online with friends, find locations and millions of information on one click of the button and help your self with getting a job with neat, prepared, printed work that your boss will love."

META FEATURES

1. Count of words.
2. Count of sentences.
3. Count of distinct words.
4. Count of syllables.

LEXICAL FEATURES

5. Count of Nouns.
6. Count of Proper nouns.
7. Count of Verbs.
8. Count of Adjectives.

DERIVED READABILITY INDICES

9. Simple Readability Index
10. Flesch Reading Ease
11. Gunning Fog Index
12. Smog Index
13. Dale Chall Readability Index

APPROACH

FEATURE GENERATION

- Generate numerical features.

"Dear @CAPS1 @CAPS2, I believe that using computers will benefit us in many ways like talking and becoming friends with others through websites like facebook and myspace. Using computers can help us find coordinates, locations, and able ourselves to millions of information. Also computers will benefit us by helping with jobs as in planning a house plan and typing a @NUM1 page report for one of our jobs in less than writing it. Now lets go into the wonder world of technology. Using a computer will help us in life by talking or making friends on line. Many people have myspace, facebooks, aim, these all benefit us by having conversations with one another. Many people believe computers are bad but how can you make friends if you can never talk to them? I am very fortunate for having a computer that can help with not only school work but my social life and how I make friends. Computers help us with finding our locations, coordinates and millions of information online. If we didn't go on the internet a lot we wouldn't know how to go onto websites that @MONTH1 help us with locations and coordinates like @LOCATION1. Would you rather use a computer or be in @LOCATION3. When your supposed to be vacationing in @LOCATION2. Million of information is found on the internet. You can ask almost every question and a computer will have it. Would you rather easily draw up a house plan on the computers or take @NUM1 hours doing one by hand with ugly eraser marks all over it, you are guaranteed that to find a job with a drawing like that. Also when applying for a job many workers must write very long papers like a @NUM3 word essay on why this job fits you the most, and many people I know don't like writing @NUM3 words non-stop for hours when it could take them I have an a computer. That is why computers we needed a lot now adays. I hope this essay has impacted your decision on computers because they are great machines to work with. The other day I showed my mom how to use a computer and she said it was the greatest invention sense sliced bread! Now go out and buy a computer to help you chat online with friends, find locations and millions of information on one click of the button and help your self with getting a job with neat, prepared, printed work that your boss will love."



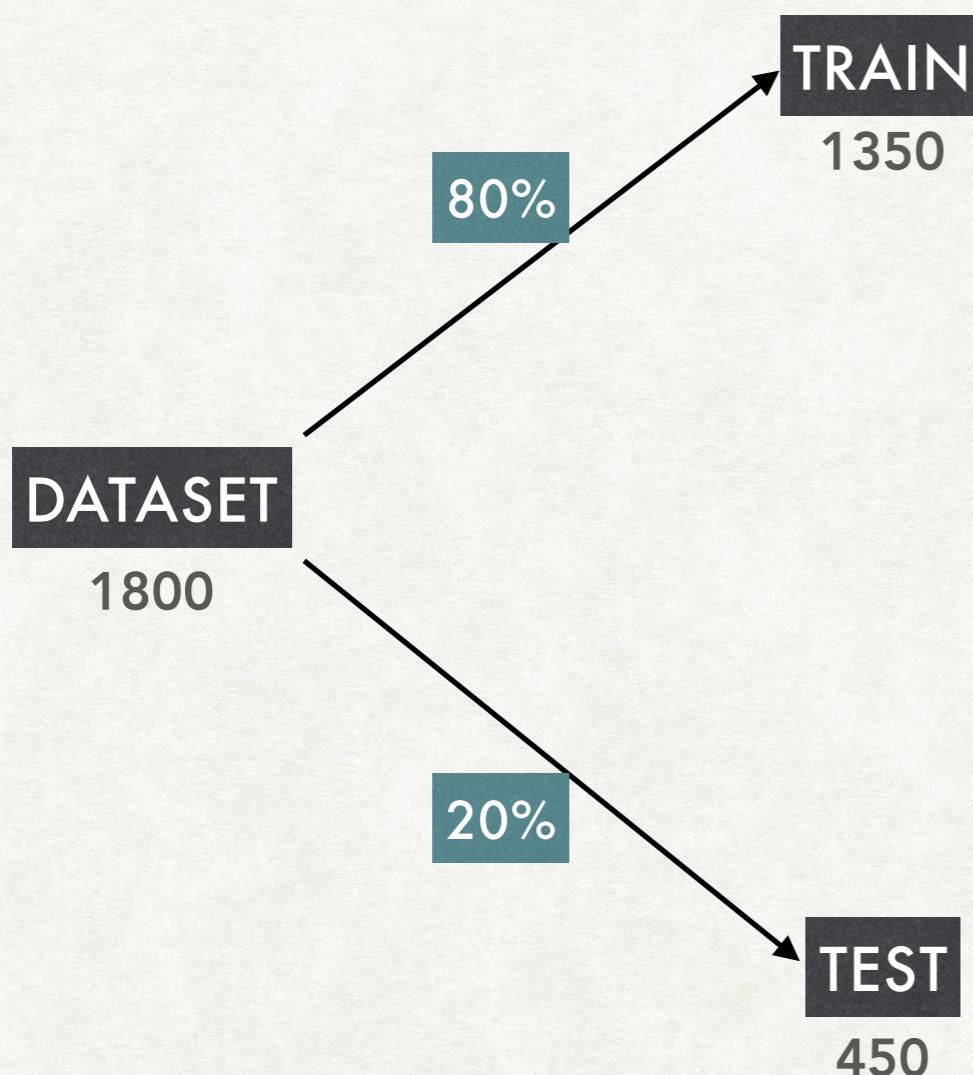
nounCount	propnCount	verbCount	adjCount	senCount	wordCount	distinctCount	syllableCount	avgSPerWord	readabilityIndex	riFRE	riGF	riSI	riDC
78	5	73	19	17	338	128	452	1.32	9.943070	74.98	10.071284	3.129	1.033

- Motivation
- Problem Statement
- Data Description
- Challenge
- Approaches
- **Results**
- Conclusions

RESULTS

PREPARING DATA

- Train/Test Split



RESULTS

LOGISTIC REGRESSION

R Squared Values

Essay Set	Train	Test
1	0.54	0.46
3	0.379	0.367
5	0.622	0.601
6	0.390	0.358

RESULTS

KNN (N=7)

R Squared Values

Essay Set	Train	Test
1	0.715	0.658
3	0.48	0.35
5	0.660	0.544
6	0.50	0.32

RESULTS

NEURAL NETWORK

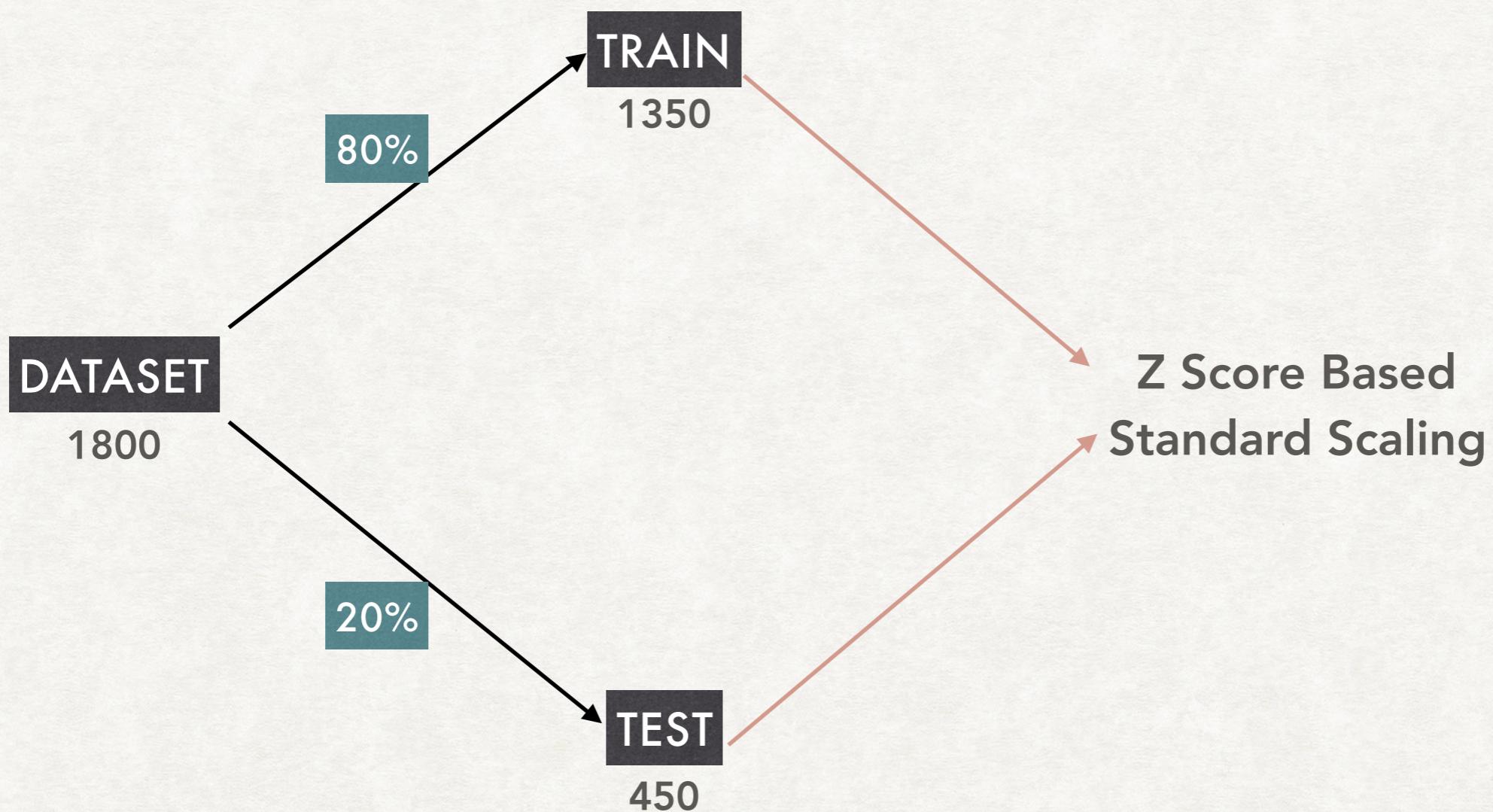
R Squared Values

Essay Set	Train	Test
1	0.78	0.72
3	0.59	0.50
5	0.74	0.67
6	0.64	0.50

RESULTS

NEURAL NETWORK

- Scaling



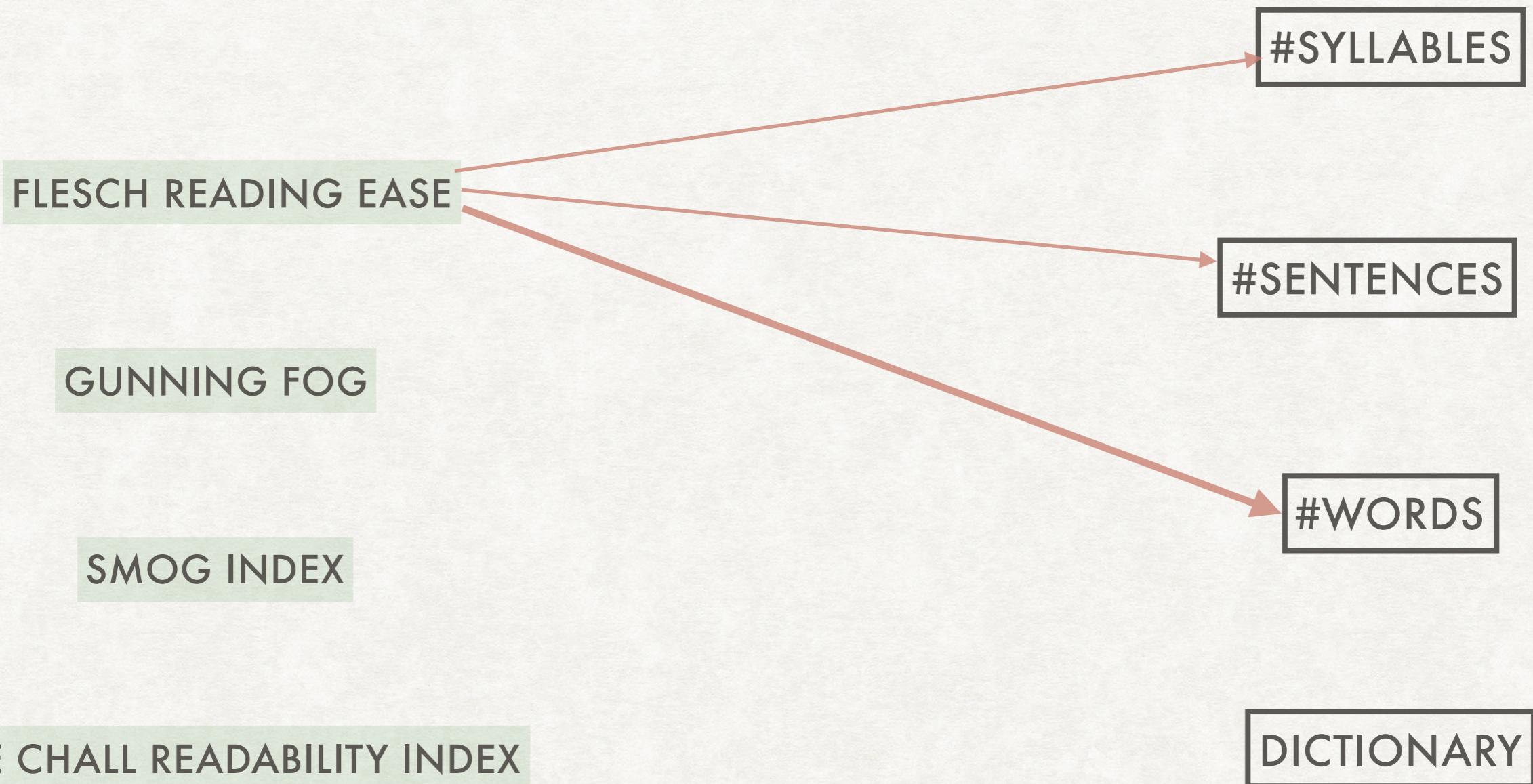
RESULTS

NEURAL NETWORK

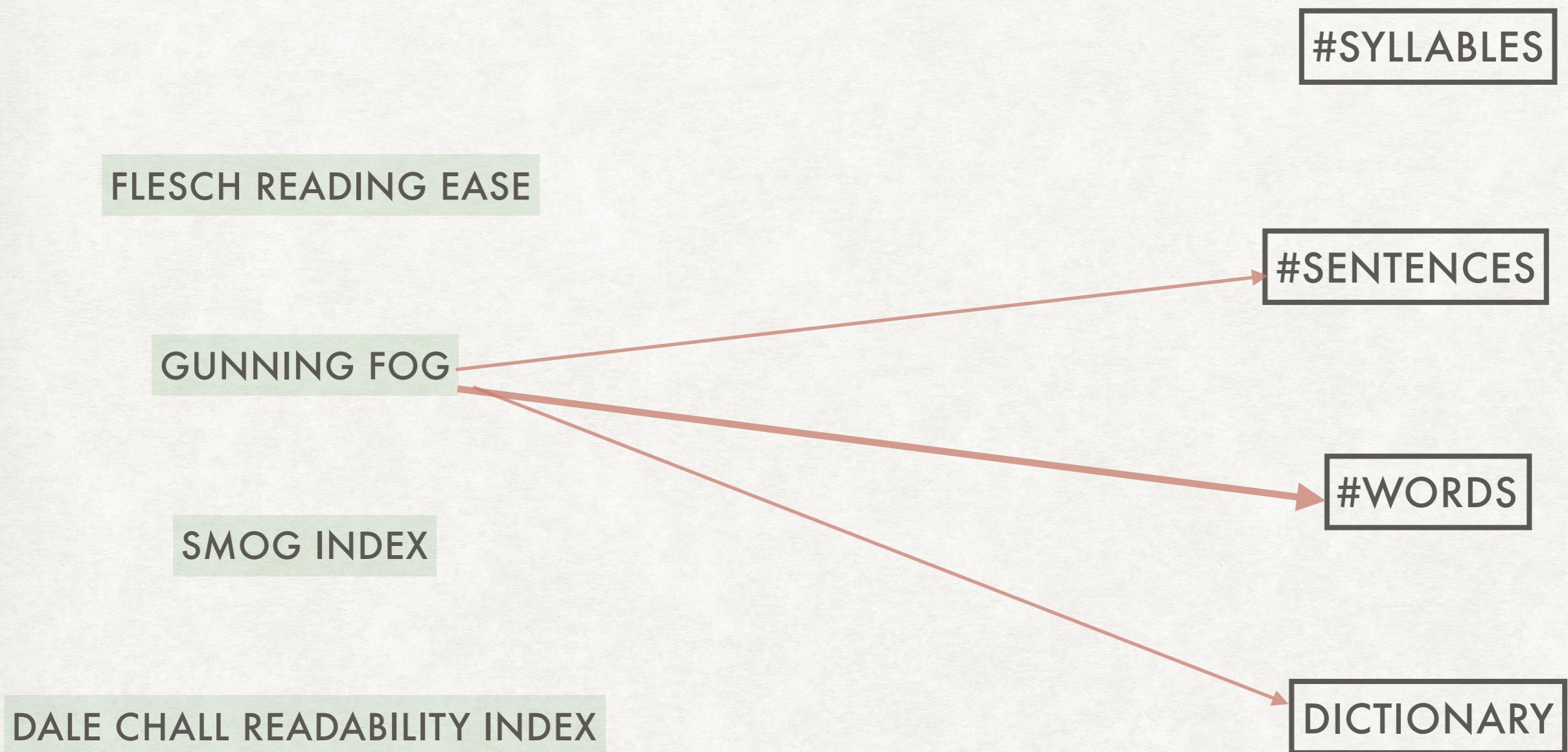
- **Input Layer - 14**
Hidden Layer 1 - 72
Hidden Layer 2 - 64
Hidden Layer 3 - 48
Output Layer - 1
- **Learning Rate - 0.001**
Optimizer - Adam
Loss Func - MSE
- **Validation Set - 0.20**
Epochs - 50

- Motivation
- Problem Statement
- Data Description
- Challenge
- Approaches
- Results
- **Summary**

SUMMARY READABILITY INDICES



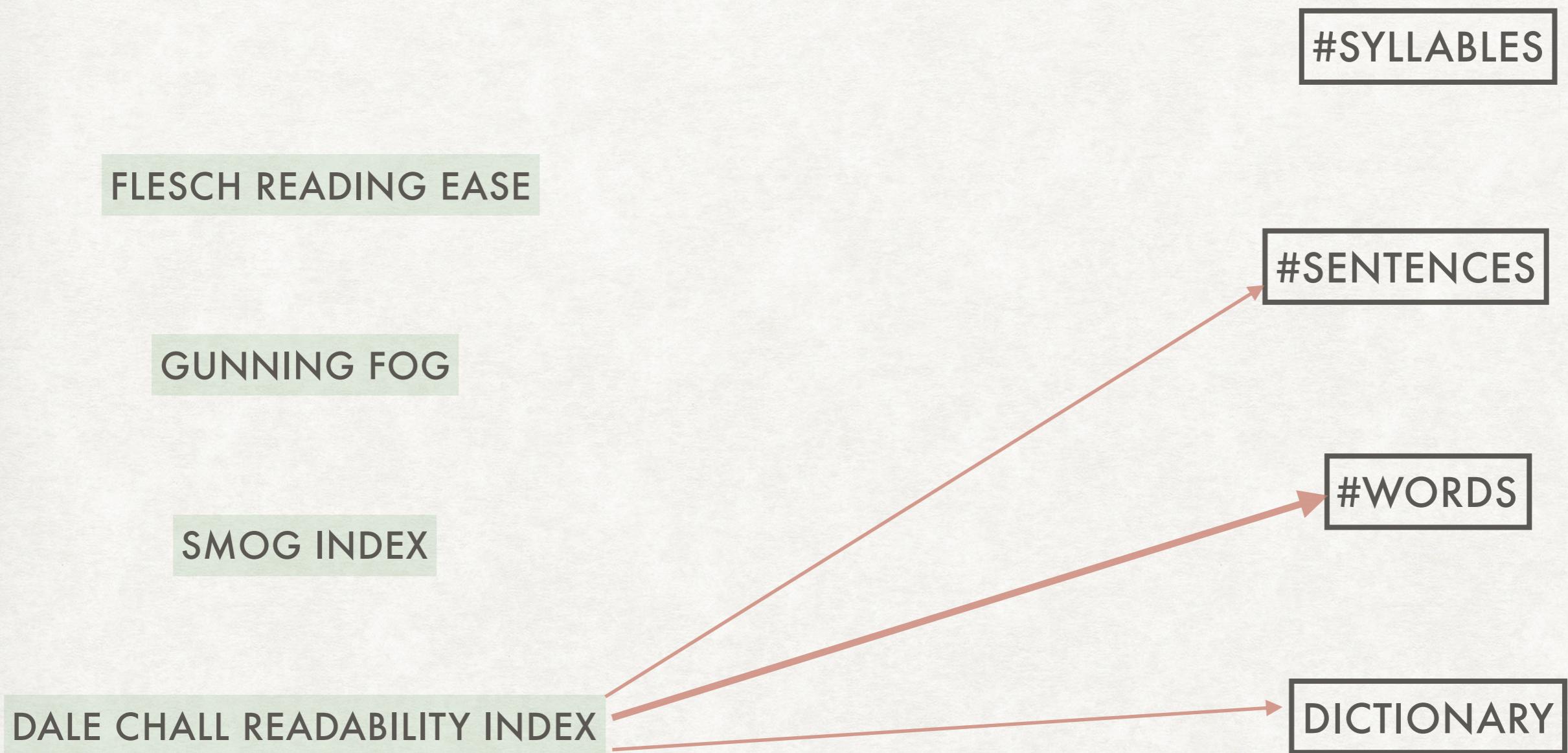
SUMMARY READABILITY INDICES



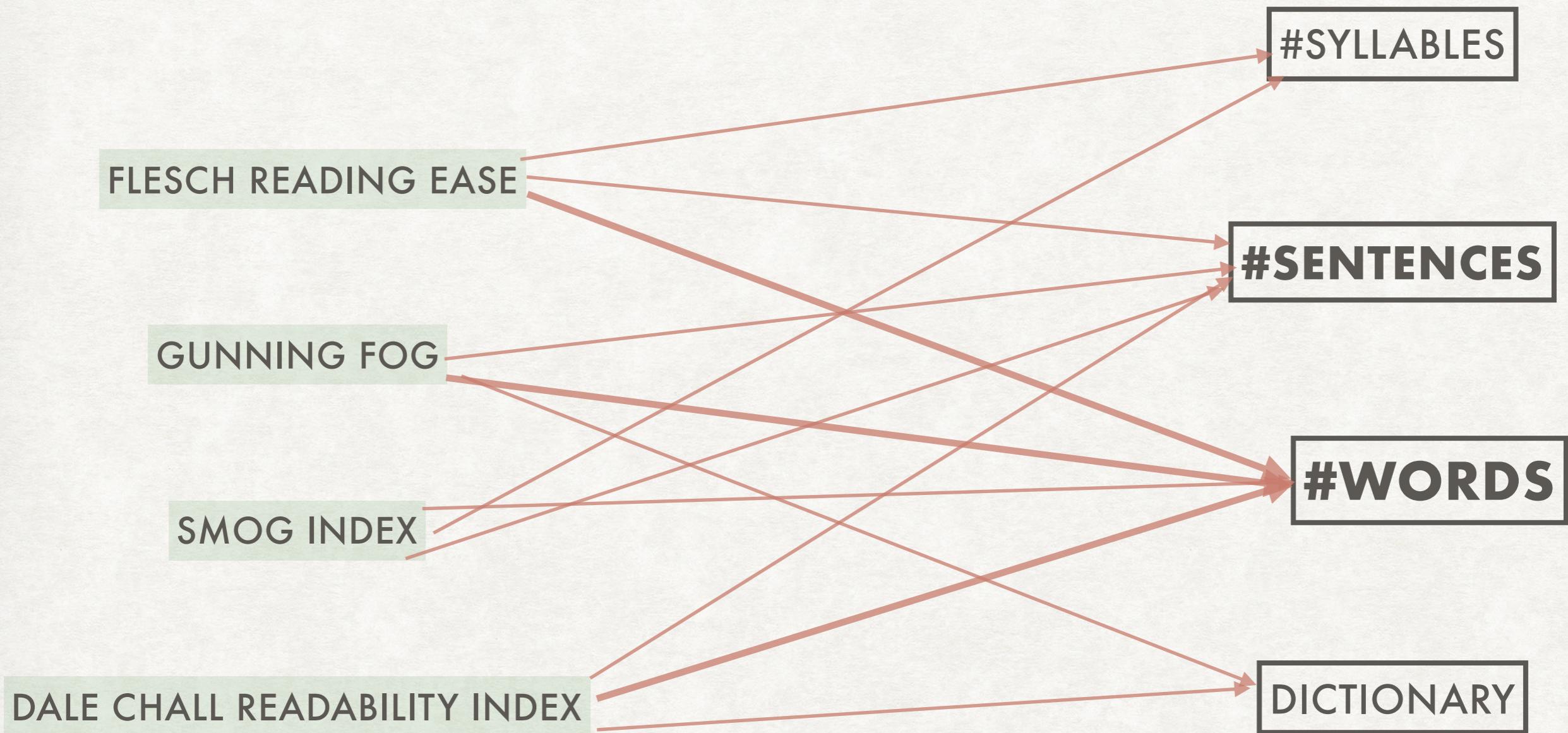
SUMMARY READABILITY INDICES



SUMMARY READABILITY INDICES



SUMMARY READABILITY INDICES



SUMMARY MODELS

- Essay set 1 showed computer predictable behavior.
- Essay sets had varying predictability.
- Models were easily overtrained.
 - Care with Learning Rate and Epochs.
- Whats ahead:
 - Factor Analysis
 - Better hyper parameter tuning?
 - Read about essays?

THANK YOU!
Q/A