國立陽明交通大學
NATIONAL YANG MING CHIAO TUNG UNIVERSITY

Digital Medicine 2021

Case Presentation I

Obesity Status Detection

2021/9/27

n2c2 NLP Research Data Sets

Unstructured notes from the Research Patient Data Repository at Partners Healthcare.

HARVARD MEDICAL SCHOOL | BLAVATNIK INSTITUTE BIOMEDICAL INFORMATICS

Need help? Contact us!

Description

The majority of these Clinical Natural Language Processing (NLP) data sets were originally created at a former NIH-funded National Center for Biomedical Computing (NCBC) known as i2b2: Informatics for Integrating Biology and the Bedside.

- 2006 - Deidentification & Smoking
- 2008 - Obesity
- 2009 - Medication
- 2010 - Relations
- 2011 - Coreference
- 2012 - Temporal Relations
- 2014 - Deidentification & Heart Disease

Based at Partners HealthCare System in Boston from 2004 to 2014, under the leadership of Principal Investigator Isaac Kohane, MD, PhD, and Executive Director Susanne Churchill, PhD, the i2b2 Center was a passionate advocate for the

- 2006 De-identification and Smoking Status Challenge Downloads
- 2008 Obesity Challenge Downloads
- 2009 Medication Challenge Downloads
- 2010 Relations Challenge Downloads
- 2011 Coreference Challenge Downloads
- 2012 Temporal Relations Challenge Downloads
- 2014 De-identification and Heart Disease Risk Factors Challenge Downloads

Electronic Medical Record

- **Discharge Date**: 3/16/1993
- **DISCHARGE DIAGNOSIS**:
  - LEFT HEMISPHERE CEREBROVASCULAR ACCIDENT SECONDARY TO OCCLUSIONS OF THE MIDDLE CEREBRAL ARTERY
- **SECONDARY DIAGNOSES**:
  - 1. ATRIAL SEPTAL DEFECT
  - 2. PROSTATIC HYPERTROPHY
  - 3. DEGENERATIVE JOINT DISEASE
  - 4. UMBILICAL HERNIA

Electronic Medical Record

- **HISTORY OF PRESENT ILLNESS**: Mr. Dorich is a 66-year-old right handed gentleman with no previous admissions to Tiftver Community , who presented on 27 of June , after having fallen out of bed at about 5:45 in the morning. He was noted to be mute at the time and was unable to move the right arm and leg. He was brought by ambulance to the Emergency Room , where much of his motor deficits resolved , as well as some of his speech deficit. This occurred within an hour. There was no loss of consciousness associated with the fall , and the patient was not noted to have any unusual movements or verbalizations , or tongue biting , or incontinence. He was "stunned" for a minute (by the report of his son) just after falling. The patient denied headaches , vertigo , double vision , hoarseness or numbness , or nausea , vomiting , or palpitations. He had had no previous neurological symptoms prior to this event. There was also no chest pain. There was no history of high blood pressure or heart disease , or recent fever and chills or sweats. He had been feeling quite well up until this point.

- **PAST MEDICAL HISTORY**: Significant for an umbilical hernia , which had been present for many years. He also suffered from prostatism. In the last year , he had regained much of the weight that he had lost previous on special diet. SOCIAL HISTORY: He did not smoke nor did he use lcohol or elicit drugs. There was no significant family history of heart disease or strokes.

- **PHYSICAL EXAMINATION**: The patient is a severely obese gentlemen, who had a blood pressure of 115/59 in the right arm; with a respiratory rate of 12; heart rate of 72. He was afebrile.

Focus on i2b2 JAMIA Obesity NLP Challenge

Viewpoint Paper ■

Recognizing Obesity and Comorbidities in Sparse Data

Özlem Uzuner, PhD

Abstract In order to survey, facilitate, and evaluate studies of medical language processing on clinical narratives, i2b2 (Informatics for Integrating Biology to the Bedside) organized its second challenge and workshop. This challenge focused on automatically extracting information on obesity and fifteen of its most common comorbidities from patient discharge summaries. For each patient, obesity and any of the comorbidities could be Present, Absent, or Questionable (i.e., possible) in the patient, or Unmentioned in the discharge summary of the patient. i2b2 provided data for, and invited the development of, automated systems that can classify obesity and its comorbidities into these four classes based on individual discharge summaries. This article refers to obesity and comorbidities as *diseases*. It refers to the categories Present, Absent, Questionable, and Unmentioned as *classes*. The task of classifying obesity and its comorbidities is called the *Obesity Challenge*.

The data released by i2b2 was annotated for textual judgments reflecting the explicitly reported information on diseases, and intuitive judgments reflecting medical professionals' reading of the information presented in discharge summaries. There were very few examples of some disease classes in the data. The Obesity Challenge paid particular attention to the performance of systems on these less well-represented classes.

A total of 30 teams participated in the Obesity Challenge. Each team was allowed to submit two sets of up to three system runs for evaluation, resulting in a total of 136 submissions. The submissions represented a combination of rule-based and machine learning approaches.

Evaluation of system runs shows that the best predictions of textual judgments come from systems that filter the potentially noisy portions of the narratives, project dictionaries of disease names onto the remaining text, apply negation extraction, and process the text through rules. Information on disease-related concepts, such as symptoms and medications, and general medical knowledge help systems infer intuitive judgments on the diseases.

■ **J Am Med Inform Assoc.** 2009;16:561–570. DOI 10.1197/jamia.M3115.

Text Annotation

- two obesity experts from the Massachusetts General Hospital Weight Center

- Two Tasks
  - Textual Task
  - Intuitive Task

- A total of 1237 discharge summaries

Textual Task

- Classify obesity as Present, Absent, Questionable, or Unmentioned based on explicitly documented information in the discharge summaries, e.g., the statement “the patient is obese”.

Intuitive Task

- Classify each disease as Present, Absent, or Questionable by applying their intuition and judgment to information in the discharge summaries, e.g., the statement “the patient weighs 230 lbs and is 5 ft 2 inches”.

General Rules of Judgement

- Textual judgments would require no reasoning.

- Intuitive judgments would generally agree with a textual Present, Absent, or Questionable judgment. The focus of the intuitive task would be on diseases marked Unmentioned.

Table 1 ■ Kappa Agreement on Textual and Intuitive Judgments

Comorbidity (Disease)	Textual Kappa	Intuitive Kappa
Asthma	0.90	0.76
Atherosclerotic CV disease (CAD)	0.78	0.81
Congestive heart failure (CHF)	0.81	0.74
Depression	0.92	0.86
Diabetes mellitus (DM)	0.91	0.87
Gallstones/cholecystectomy	0.92	0.90
GERD	0.89	0.59
Gout	0.93	0.92
Hypercholesterolemia	0.87	0.68
Hypertension (HTN)	0.82	0.67
Hypertriglyceridemia	0.71	0.72
Obesity	0.91	0.86
Obstructive sleep apnea (OSA)	0.92	0.92
Osteoarthritis (OA)	0.76	0.76
Peripheral vascular disease (PVD)	0.94	0.73
Venous insufficiency	0.79	0.44

CV = cardiovascular; GERD = gastroesophageal reflux disease.

Table 3 ■ Distribution of Textual Judgments into Training and Test Sets

Diseases	Present		Absent		Questionable		Unmentioned		Total	
	Training	Test	Training	Test	Training	Test	Training	Test	Training	Test
Asthma	93	68	3	2	2	2	630	432	728	504
CAD	399	277	23	22	7	2	292	196	721	497
CHF	310	205	11	11	0	0	399	280	720	496
Depression	104	72	0	0	0	0	624	434	728	506
DM	485	338	15	12	7	3	219	150	726	503
Gallstones	109	87	4	2	1	0	615	418	729	507
GERD	118	69	1	1	5	1	599	433	723	504
Gout	90	52	0	0	4	0	634	453	728	505
Hypercholesterolemia	304	213	13	6	1	4	408	279	726	502
HTN	537	374	12	6	0	0	180	121	729	501
Hypertriglyceridemia	18	10	0	0	0	0	711	497	729	507
Obesity	298	198	4	3	4	3	424	289	730	493
OSA	105	69	1	0	8	2	614	432	728	503
OA	115	86	0	0	0	0	613	416	728	502
PVD	102	64	0	0	0	0	627	497	729	507
Venous insufficiency	21	10	0	0	0	0	707	497	728	507
Total	3208	2192	87	65	39	17	8296	5770	11630	8044

CAD = coronary artery disease; CHF = congestive heart failure; DM = diabetes mellitus; GERD = gastroesophageal reflux disease; HTN = hypertension; OSA = obstructive sleep apnea; OA = osteo arthritis; PVD = peripheral vascular disease.

Table 4 ■ Distribution of Intuitive Judgments into Training and Test Sets

Diseases	Present		Absent		Questionable		Total	
	Training	Test	Training	Test	Training	Test	Training	Test
Asthma	86	68	596	403	0	0	682	471
CAD	391	272	265	185	5	1	661	458
CHF	308	205	318	229	1	4	627	438
Depression	142	105	555	372	0	0	697	477
DM	473	333	205	146	5	0	683	479
Gallstones	101	80	609	411	0	0	710	491
GERD	144	93	447	331	1	2	592	426
Gout	94	61	616	439	2	0	712	500
Hypercholesterolemia	315	242	287	189	1	0	603	431
HTN	511	358	127	88	0	0	638	446
Hypertriglyceridemia	37	25	665	461	0	0	702	486
Obesity	285	192	379	255	1	0	665	447
OSA	99	66	606	427	8	2	713	495
OA	117	91	554	367	1	4	672	462
PVD	110	65	556	399	1	1	667	465
Venous insufficiency	54	29	577	398	0	0	631	427
Total	3267	2285	7362	5100	26	14	10655	7399

CAD = coronary artery disease; CHF = congestive heart failure; DM = diabetes mellitus; GERD = gastroesophageal reflux disease; HTN = hypertension; OSA = obstructive sleep apnea; OA = osteo arthritis; PVD = peripheral vascular disease.

Table 7 ■ Micro- and Macro-averaged Results on Textual Judgments, Sorted by Macro-averaged F-Measure

Systems	Macro-Averaged			Micro-Averaged		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure
Yang et al.	0.8482	0.7737	**0.8052**	0.9723	0.9723	0.9723
Solt et al.	0.8318	0.7776	0.8000	0.9756	0.9756	0.9756
Ware et al.	0.8314	0.7542	0.7821	0.9718	0.9718	0.9718
Childs et al.	0.8169	0.7454	0.7762	0.9773	0.9773	**0.9773**
Mishra et al.	0.7485	0.8050	0.7718	0.9704	0.9704	0.9704
Szarvas et al.	0.7644	0.7600	0.7622	0.9729	0.9729	0.9729
Savova et al.	0.7701	0.7147	0.7377	0.9668	0.9668	0.9668
Patrick et al.	0.7971	0.6219	0.6737	0.9693	0.9693	0.9693
*Jazayeri et al.	0.7849	0.5779	0.6205	0.9514	0.9514	0.9514
†DeShazo et al.	0.8552	0.6240	0.6140	0.9639	0.9639	0.9639

Best F-measures are in bold. †System utilized external annotators.
*System description not available.

Table 9 ■ Micro- and Macro-averaged Results on Intuitive Judgments, Sorted by Macro-averaged F-Measure

Systems	Macro-Averaged			Micro-Averaged		
	Precision	Recall	F-Measure	Precision	Recall	F-Measure
Solt et al.	0.7485	0.6571	**0.6745**	0.9590	0.9590	0.9590
Szarvas et al.	0.6999	0.6588	0.6727	0.9642	0.9642	0.9642
Childs et al.	0.7061	0.6540	0.6696	0.9582	0.9582	0.9582
Ware et al.	0.6410	0.6399	0.6404	0.9654	0.9654	**0.9654**
Ambert et al.	0.6383	0.6307	0.6344	0.9558	0.9558	0.9558
Meystre	0.6304	0.6387	0.6343	0.9566	0.9566	0.9566
Yang et al.	0.6383	0.6294	0.6336	0.9572	0.9572	0.9572
†DeShazo et al.	0.9722	0.6216	0.6292	0.9524	0.9523	0.9524
Matthews	0.6325	0.6256	0.6288	0.9509	0.9509	0.9509
Jazayeri et al.	0.6320	0.6257	0.6287	0.9508	0.9508	0.9508

Best F-measures are in bold. †System utilized external annotators.

Case Presentation 1

- **Design and present** an analysis flow for obesity status classifiers according to textual judgement (**presence of obesity or unmentioned**).

- Training data based on textual judgement
  - Textual judgement: 200 cases obesity vs. 200 cases unmentioned.

- Testing data based on intuitive judgement
  - Intuitive judgement: 200 cases obesity vs. 200 cases absence

- Validation data (50 cases) based on textual judgement

Training Data

名稱	修改日期	類型	大小
U_ID_865.txt	2021/9/26 上午 10:31	文字文件	5 KB
U_ID_866.txt	2021/9/26 上午 10:31	文字文件	6 KB
U_ID_867.txt	2021/9/26 上午 10:31	文字文件	8 KB
U_ID_869.txt	2021/9/26 上午 10:31	文字文件	4 KB
U_ID_871.txt	2021/9/26 上午 10:31	文字文件	8 KB
U_ID_872.txt	2021/9/26 上午 10:31	文字文件	6 KB
U_ID_874.txt	2021/9/26 上午 10:31	文字文件	5 KB
U_ID_877.txt	2021/9/26 上午 10:31	文字文件	2 KB
U_ID_878.txt	2021/9/26 上午 10:31	文字文件	8 KB
U_ID_879.txt	2021/9/26 上午 10:31	文字文件	10 KB
U_ID_882.txt	2021/9/26 上午 10:31	文字文件	6 KB
U_ID_884.txt	2021/9/26 上午 10:31	文字文件	3 KB
U_ID_885.txt	2021/9/26 上午 10:31	文字文件	7 KB
U_ID_889.txt	2021/9/26 上午 10:31	文字文件	5 KB
U_ID_891.txt	2021/9/26 上午 10:31	文字文件	11 KB
Y_ID_6.txt	2021/9/26 上午 10:31	文字文件	7 KB
Y_ID_13.txt	2021/9/26 上午 10:31	文字文件	16 KB
Y_ID_20.txt	2021/9/26 上午 10:31	文字文件	9 KB
Y_ID_21.txt	2021/9/26 上午 10:31	文字文件	12 KB
Y_ID_28.txt	2021/9/26 上午 10:31	文字文件	4 KB
Y_ID_50.txt	2021/9/26 上午 10:31	文字文件	4 KB
Y_ID_51.txt	2021/9/26 上午 10:31	文字文件	3 KB
Y_ID_56.txt	2021/9/26 上午 10:31	文字文件	4 KB
Y_ID_62.txt	2021/9/26 上午 10:31	文字文件	6 KB
Y_ID_65.txt	2021/9/26 上午 10:31	文字文件	6 KB

Testing Data

N_ID_929	2021/9/26 上午 10:35	文字文件	6 KB
N_ID_932	2021/9/26 上午 10:35	文字文件	7 KB
N_ID_934	2021/9/26 上午 10:35	文字文件	11 KB
N_ID_935	2021/9/26 上午 10:35	文字文件	4 KB
N_ID_939	2021/9/26 上午 10:35	文字文件	8 KB
N_ID_940	2021/9/26 上午 10:35	文字文件	5 KB
N_ID_943	2021/9/26 上午 10:35	文字文件	8 KB
N_ID_945	2021/9/26 上午 10:35	文字文件	12 KB
N_ID_948	2021/9/26 上午 10:35	文字文件	8 KB
Y_ID_13	2021/9/26 上午 10:35	文字文件	16 KB
Y_ID_28	2021/9/26 上午 10:35	文字文件	4 KB
Y_ID_50	2021/9/26 上午 10:35	文字文件	4 KB
Y_ID_51	2021/9/26 上午 10:35	文字文件	3 KB
Y_ID_56	2021/9/26 上午 10:35	文字文件	4 KB
Y_ID_65	2021/9/26 上午 10:35	文字文件	6 KB
Y_ID_68	2021/9/26 上午 10:35	文字文件	13 KB
Y_ID_70	2021/9/26 上午 10:35	文字文件	4 KB
Y_ID_84	2021/9/26 上午 10:35	文字文件	6 KB

Validation Data

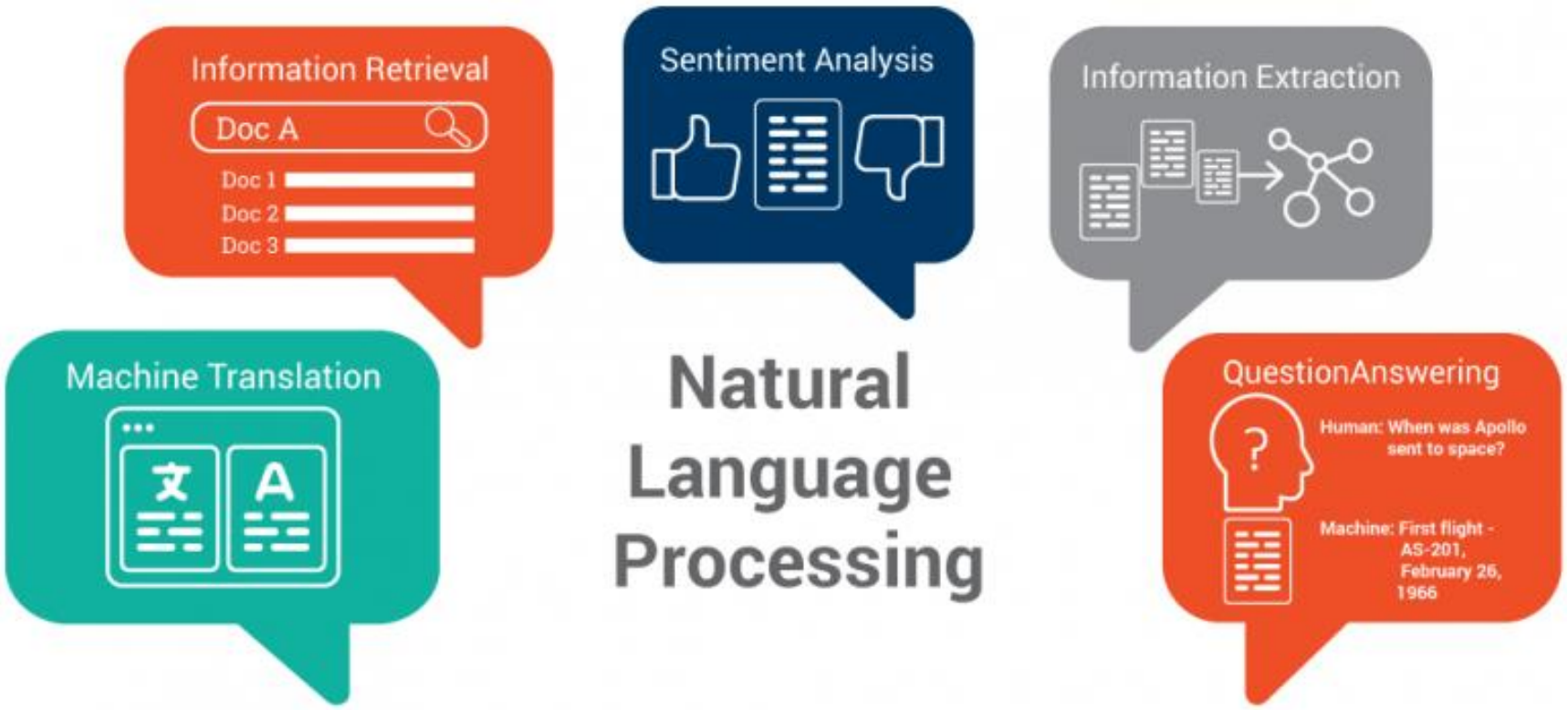
ID_1159	2021/9/26 上午 10:31	文字文件	7 KB
ID_1160	2021/9/26 上午 10:31	文字文件	7 KB
ID_1162	2021/9/26 上午 10:31	文字文件	5 KB
ID_1167	2021/9/26 上午 10:31	文字文件	8 KB
ID_1168	2021/9/26 上午 10:31	文字文件	8 KB
ID_1176	2021/9/26 上午 10:31	文字文件	6 KB
ID_1180	2021/9/26 上午 10:31	文字文件	9 KB
ID_1183	2021/9/26 上午 10:31	文字文件	4 KB
ID_1184	2021/9/26 上午 10:31	文字文件	11 KB
ID_1185	2021/9/26 上午 10:31	文字文件	4 KB
ID_1186	2021/9/26 上午 10:31	文字文件	6 KB
ID_1187	2021/9/26 上午 10:31	文字文件	4 KB
ID_1189	2021/9/26 上午 10:31	文字文件	8 KB
ID_1190	2021/9/26 上午 10:31	文字文件	6 KB
ID_1191	2021/9/26 上午 10:31	文字文件	14 KB
ID_1193	2021/9/26 上午 10:31	文字文件	6 KB

Presentation Scoring Method

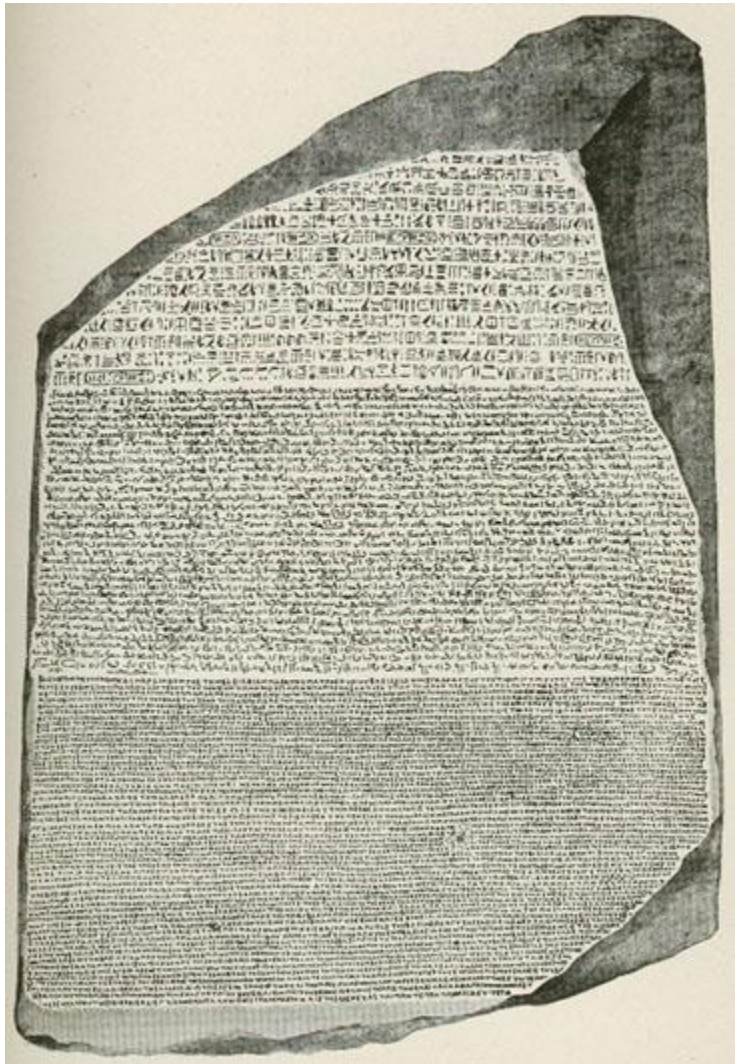
- Time Control: 10 points (10 minutes)
- Background: 10 points
- Methods (Analysis Flow): 30 points
- Validation Data Results: 20 points
- Discussion: 30 points

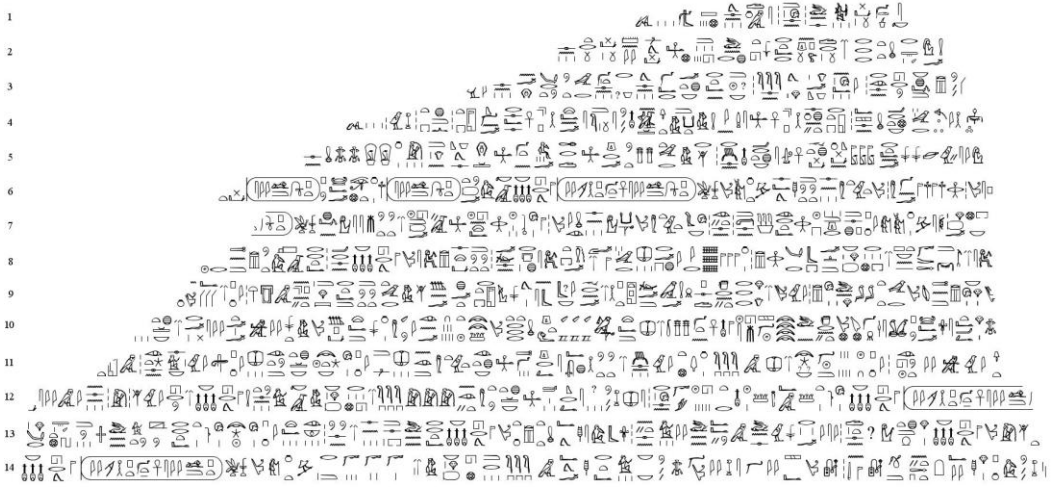
Mini-workshop of Case Presentation 1

Natural Language Processing

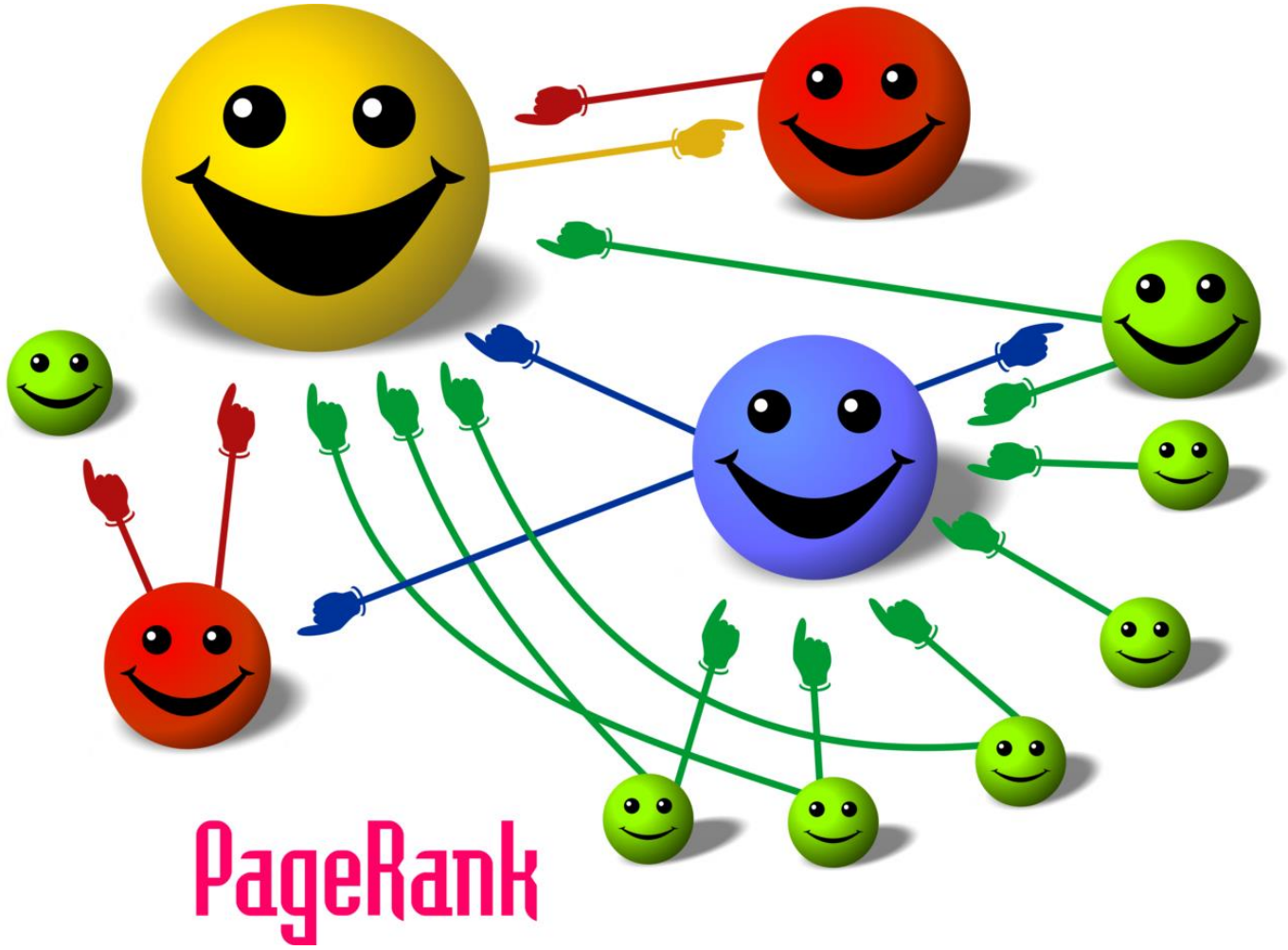


Translation

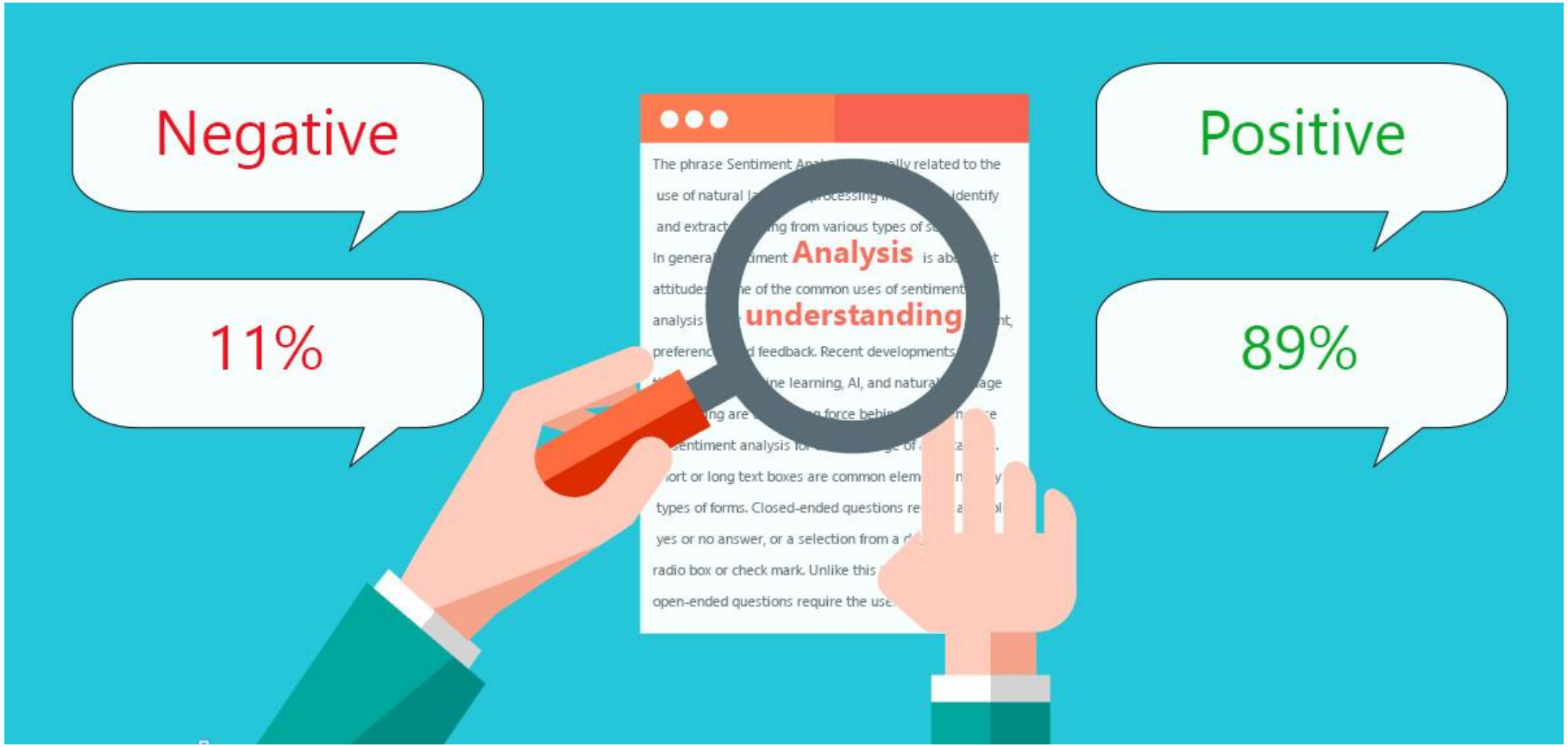




Information Retrieval



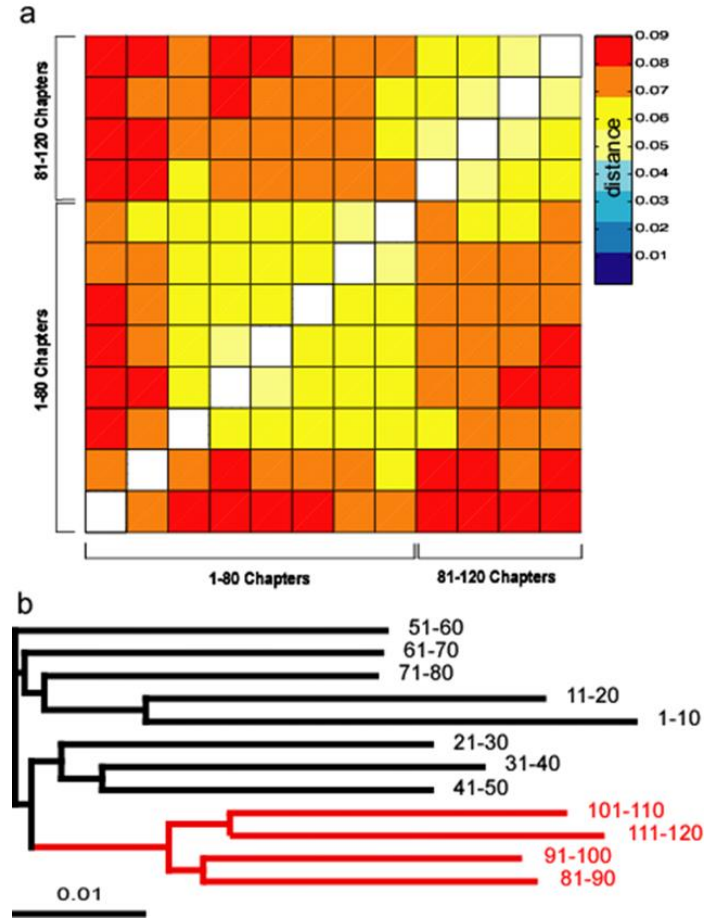
Sentiment Analysis



Information Extraction

紅樓夢

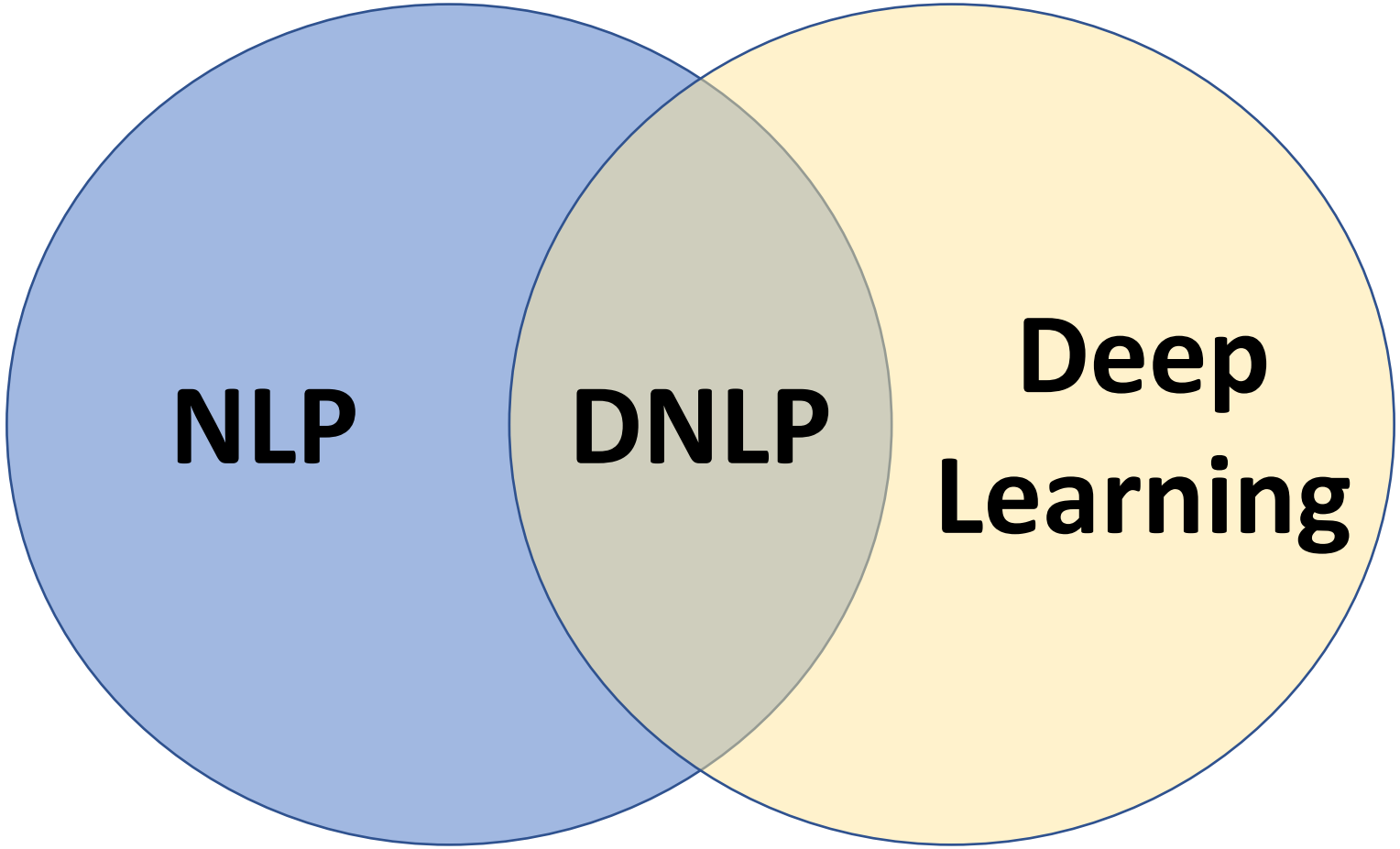
Rank	1-40		41-80		81-120	
	Word	Frequency	Word	Frequency	Word	Frequency
1	了	6250	了	8301	了	6946
2	不	4505	不	5676	的	5499
3	的	4010	的	5539	不	5009
4	一	3891	一	4942	來	3944
5	道	3683	來	4097	道	3756
6	來	3563	人	3892	是	3741
7	人	3139	我	3769	人	3644
8	我	2843	是	3720	一	3461
9	是	2833	道	3683	說	3391
10	說	2805	說	3637	我	2743



Physical Review Letters 90:108103 (2003). **Physica A** 329:473-483 (2003)

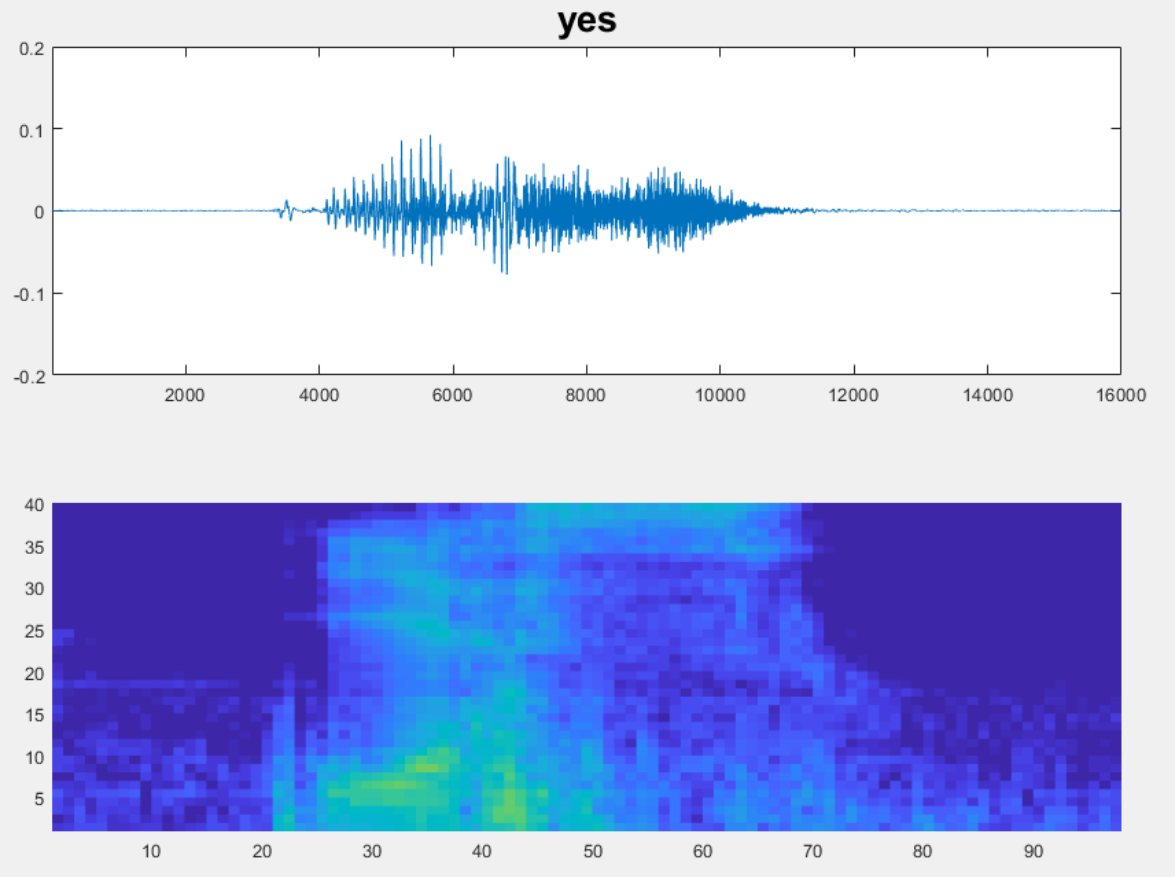
Question Answering

Types of Natural Language Processing (NLP)



Classical NLP Models

- Audio frequency component analysis (Speech Recognition)



Classical NLP Models

- If/else rules (ChatBot & Classification)

Classical NLP Models

- If/else rules for obesity status detection

```
obese = {‘obese’, ‘obesity’, ‘overweight’}
If presence(‘obese’)
   obese_status = ‘yes’
else
    obese_status = ‘unmentioned’
end
```

Classical NLP Models

- Bag-of-words model (classification)

Comment	Pass/Fail
Great job!	1
Amazing work.	1
Well done.	1
Very well written.	1
Poor effort.	0
Could have done better.	0
Try harder next time.	0
...	...

Bag-of-Words Model

List of Words	Frequency of Words
obese	0
obesity	1
overweight	0
weight gain	1
…	…

Convolutional Neural Network for Text Classification

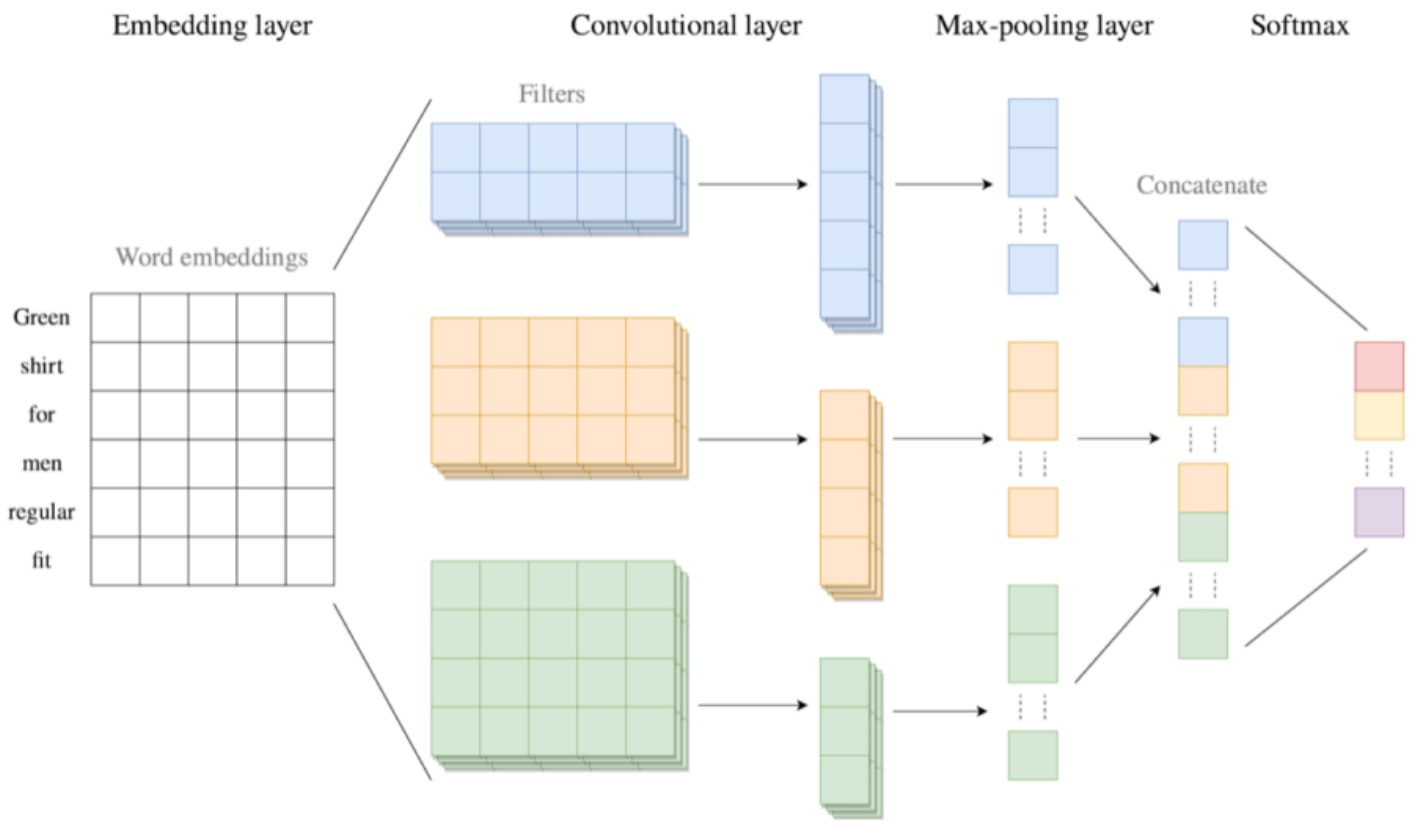
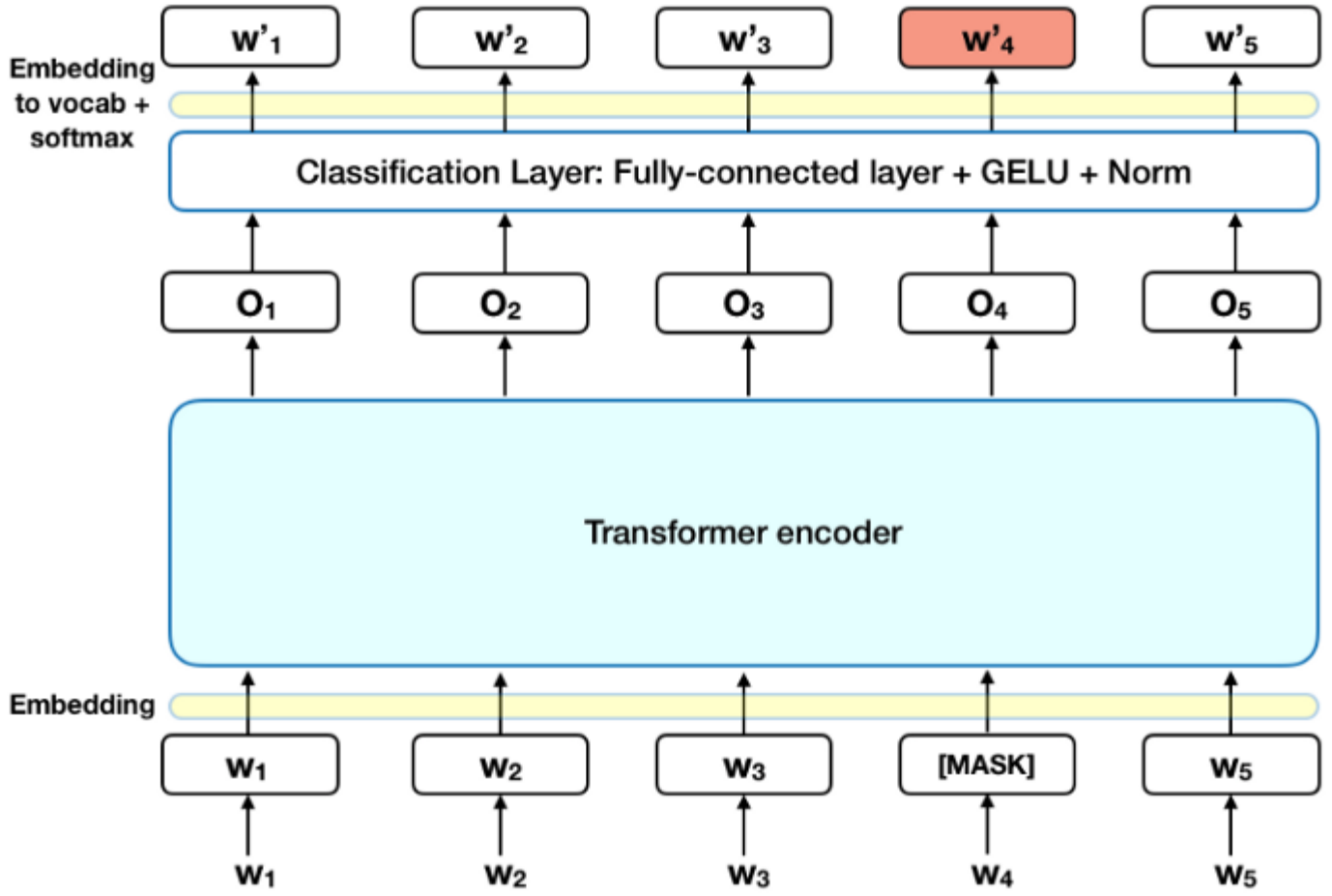


Figure 3: General CNN architecture for text classification

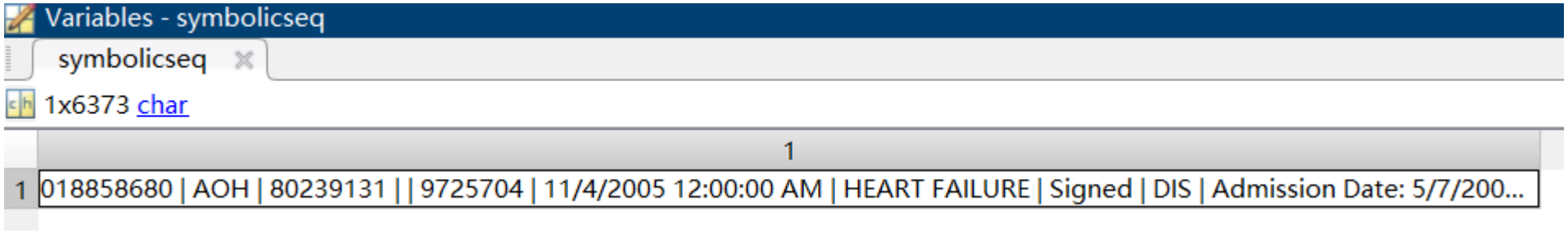
BERT
(Bidirectional Encoder Representations from Transformers)



https://towardsdatascience.com/bert-explained-state-of-the-art-language-model-for-nlp-f8b21a9b6270

Read Text File Into Matlab

- symbolicseq = fileread('Y_ID_6.txt');



Y_ID_6 - 記事本

檔案(F) 編輯(E) 格式(O) 檢視(V) 說明

018858680 | AOH | 80239131 | | 9725704 | 11/4/2005 12:00:00 AM | HEART FAILURE | Signed | DIS | Admission Date: 5/7/2005 Report Status: Signed

Discharge Date: 10/8/2005
ATTENDING: SATOW , DARNELL M.D.
CHIEF COMPLAINT:
This is a 64-year-old woman with a history of nonischemic
cardiomyopathy who was admitted with increasing volume overload.
HISTORY OF PRESENT ILLNESS:
Ms. Schmittou was doing well until approximately two weeks prior to
admission when in the setting of dietary indiscretion she noticed
some shortness of breath and weight gain. She states that she
started to increase her Demadex from 200 mg in the morning and
150 mg in the afternoon to 200 mg b.i.d. without significant
diuresis. She has continued to participate in cardiac rehab
without significant difficulty. Her weight was stable at 174.
She did , however , noticed worsening of intermittent dyspnea on
exertion and orthopnea , although she denied PND or pedal edema.
She denied abdominal symptoms of bloating , no satiety , nausea or
vomiting. She denied chest pain , discomfort , palpitations ,
lightheadedness or syncope. She does have occasional chest
burning. She also noted burning in her feet bilaterally which
has previously been attributed to neuropathy. On the day of
admission , she presented to the Outpatient Clinic to have an
echocardiogram and repeat exercise stress test with oxygen
uptake. She exercised for 4 minutes and 32 seconds achieving a
peak heart of 90 beats per minute with a blood pressure of 100/64
and stopping secondary to fatigue. Her peak oxygen uptake was
only 9 mL per minute despite on RER of 1.09. This is essentially
unchanged from a prior exercise stress test one year ago.
PAST MEDICAL HISTORY:
Includes nonischemic cardiomyopathy diagnosed in 1996 , with clean
coronary catheterization at that point. She has history of

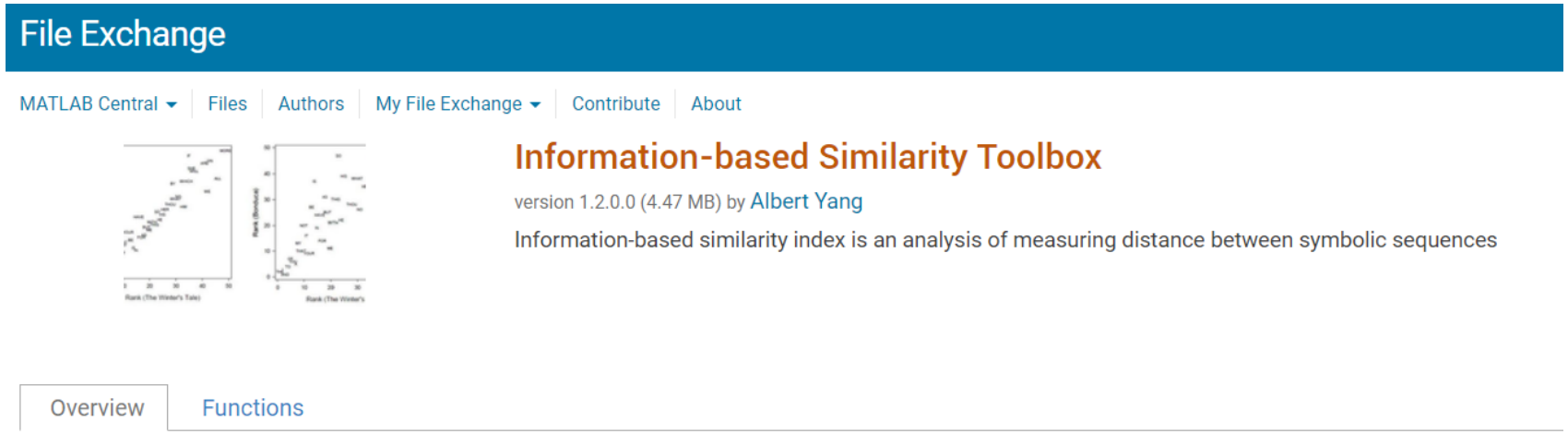
Preprocessing Text

- symbolicseq = text_preprocessing(symbolicseq,0);

symbolicseq

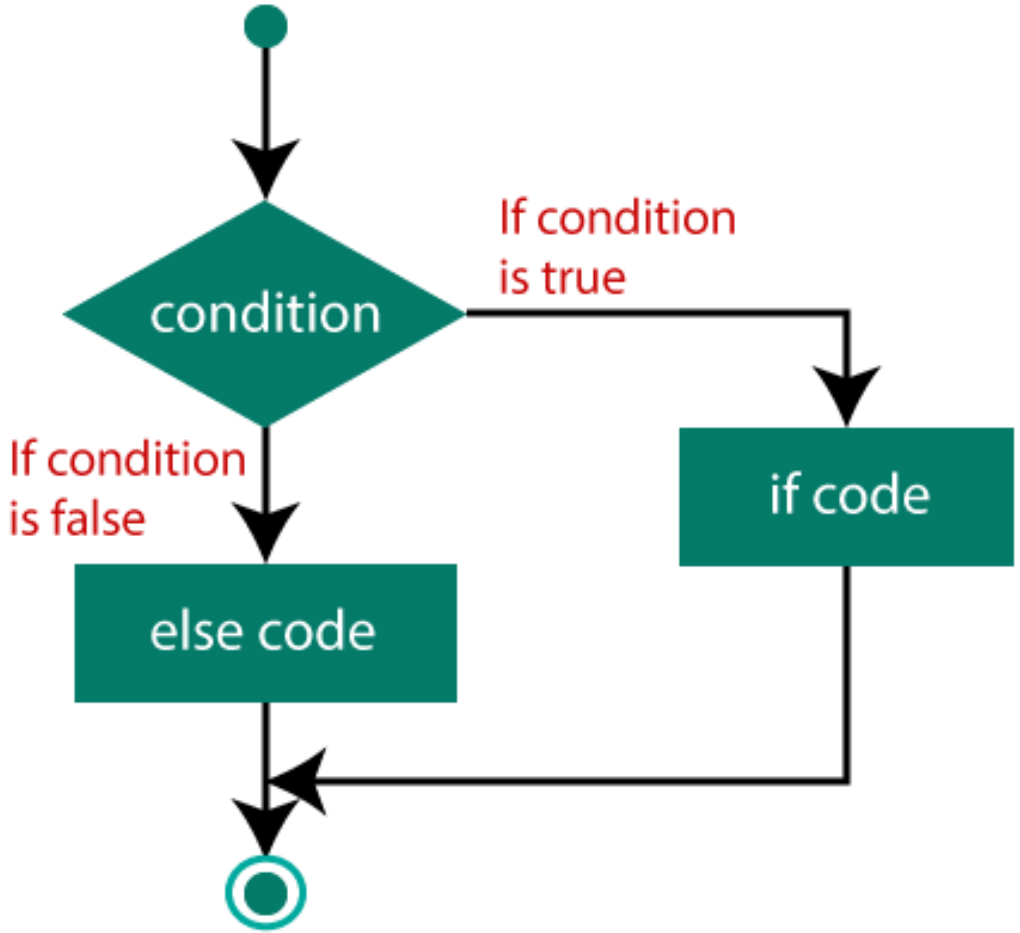
295x1 cell

	1	2
1	YC	
2	AM	
3	OBSTRUC...	
4	SLEEP	
5	APNEA	
6	DIS	
7	ADMISSI...	
8	DATE	
9	REPORT	
10	STATUS	
11	DISCHAR...	
12	DATE	
13	DISCHAR...	
14	ORDERS	
15	BREUNCL...	
16	BINTRIREE	



https://www.mathworks.com/matlabcentral/fileexchange/46691-information-based-similarity-toolbox

If / Else Rule


condition
If condition
is true
if code
If condition
is false
else code

List of Words Related to Obesity

- obese
- obesity
- overweight
- …

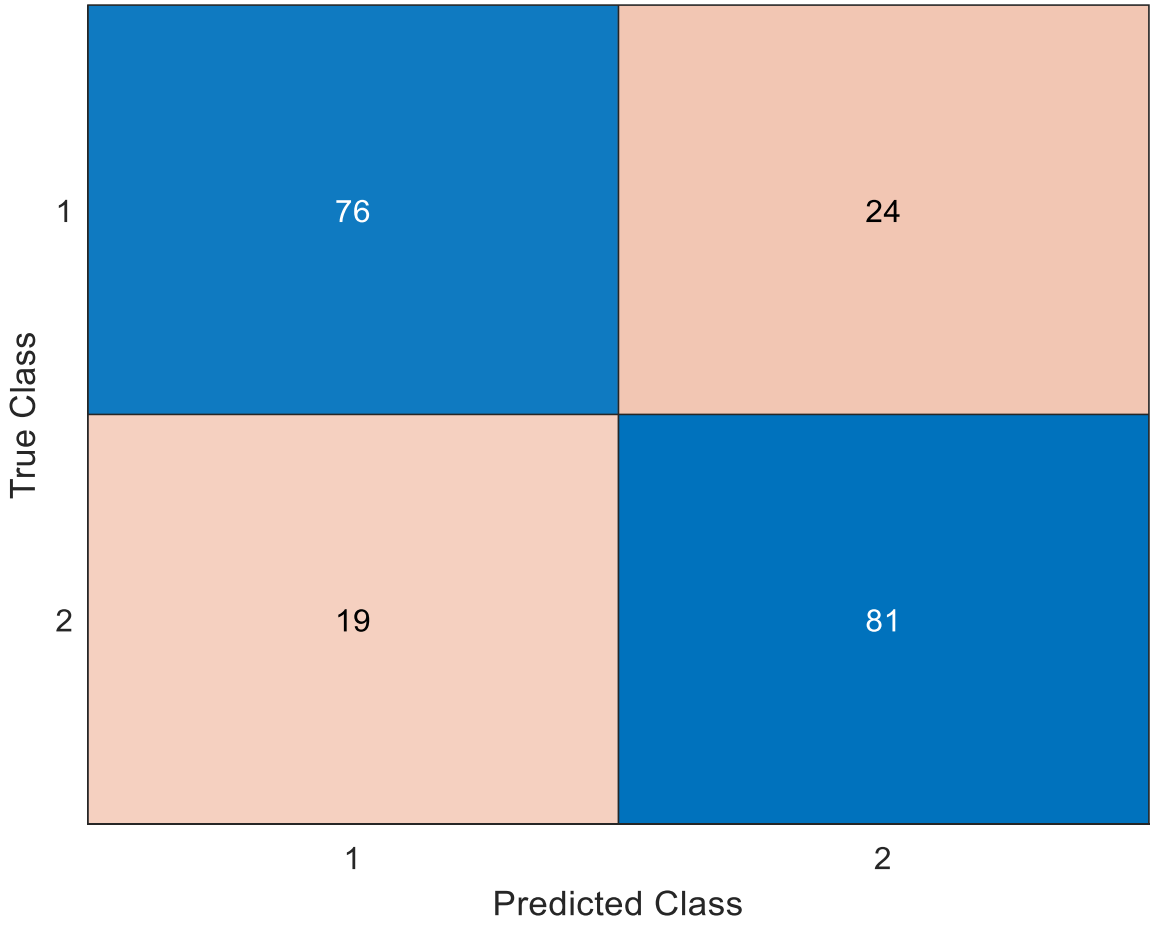
If / Else Rule to Identify Obesity Keywords

```
obese_words = {'obese','obesity'};
obese_counts = 0;
For i=1:length(symbolicseq)
    match = cellfun(@(s) ~isempty(strfind(symbolicseq{i}, s)), obese_words)
    if sum(match) > 0
        obese_counts = obese_counts + 1;
    end
end
```

If / Else Rule to Identify Obesity Status

```
if obese_counts > 0
    obesity = true
else
    obesity = false
end
```

Confusion Matrix



Confusion Matrix Statistics

76	24
19	81

Observed Class		Predicted Class: No	Predicted Class: Yes
	No	TN	FP
	Yes	FN	TP

TN	True Negative
FP	False Positive
FN	False Negative
TP	True Positive

Model Performance

(76+81)/200 = 0.785

Accuracy	= (TN+TP)/(TN+FP+FN+TP)	
Precision	= TP/(FP+TP)	81/(24+81) = 0.77
Sensitivity	= TP/(TP+FN)	81/(19+81) = 0.81
Specificity	= TN/(TN+FP)	76/(76+24) = 0.76

Binary Confusion Matrix

https://builtin.com/data-science/supervised-machine-learning-classification

Confusion Matrix Statistics

81	24
19	76

CONFUSUSION MATRIX	ACTUAL	
PREDICTED	True Positive (TP)	False Positive (FP)
	False Negative (FN)	True Negative (TN)

81/(24+81) = 0.77

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN}$$

$$\text{F1-Score} = \frac{2*Precision*Recall}{Precision+Recall}$$

$$\text{Recall} = \frac{TP}{TP+FN}$$

2*0.77*0.81 / (0.77+0.81) = 0.79

81/(81+19) = 0.81