Copyright Notice

These slides are distributed under the Creative Commons License. DeepLearning.AI makes these slides available for educational purposes. You may not use or distribute these slides for commercial purposes. You may make copies of these slides and use or distribute them for educational purposes as long as you cite DeepLearning.AI as the source of the slides. For the rest of the details of the license, see https://creativecommons.org/licenses/by-sa/2.0/legalcode

Week 2

AI for Good Framework

W2 Lesson 1

AI for Good Framework

DeepLearning.AI

AI and Public Health

AI for Good Framework

AI for Good framework

1. Engage stakeholders
2. Define the problem
3. Determine if AI could add value

1. Prototype your solution
2. Ensure data privacy
3. Design the user experience

1. Productionize AI models
2. Integrate the user experience
3. Test with end users

1. Measure project impact
2. Communicate results
3. Determine next steps

DeepLearning.AI

AI and Public Health

Maternal and Infant Health in Nigeria

Evolution of cell phones (worldwide)

Data from The World Bank Data, CC BY-4.0, 2023

Malawi

Malawi clinic

1
physician

>100,000
residents

People communicated through text messages

Wikimedia Commons, 2016. Malawi location

Text message classification at Malawi

Input → **Output**

Text Messages

Administrative related to the clinic →
1. Patient-related
2. Clinic-admin
3. Technological

Requests from Community Health Workers →
1. Response
2. Request for doctor
3. Medical advice

Illness monitoring bodies →
1. TB
2. HIV
3. Death

Text message classification

Subword Variation in Text Message Classification

Robert Munro
Department of Linguistics
Stanford University
Stanford, CA 94305
rmunro@stanford.edu

Christopher D. Manning
Department of Computer Science
Stanford University
Stanford, CA 94305
manning@stanford.edu

Abstract

For millions of people in less resourced regions of the world, text messages (SMS) provide the only regular contact with their doctor. Classifying messages by medical labels supports rapid responses to emergencies, the early identification of epidemics and everyday administration, but challenges include text-brevity, rich morphology, phonological variation, and limited training data. We present

with costs making texts the dominant communication method. This has led social development organizations to leverage mobile technologies to support health (Leach-Lemens, 2009), banking (Peevers et al., 2008), access to market information (Jagun et al., 2008), literacy (Isbrandt, 2009) and emergency response (Munro, 2010). The possibility to automate many of these services through text-classification is huge, as are the potential benefits – those with the least resources have the most to gain.

[Munro & Manning. (2010). Subword Variation in Text Message Classification. *NAACL*]

An earthquake struck Haiti on January 12, 2010
Most cell-towers remained functional, while most local services failed

People were requesting help, mainly in Haitian Kreyol

Input → **Output**

Fanmi mwen nan Kafou, 24 Cote Plage, 4 I A bezwen manje ak dlo

My family in Carrefour, 24 Cote Plage, 4 I A needs food and water

Moun kwense nan Sakre Kè nan Pòtoprens

People trapped in Sacred Heart Church, PauP

2000 volunteers globally, transferred to paid workers in Haiti

Mission 4636

“Fanm gen tranche pou fè yon pitit nan Delmas 31”

Message translated, categorized & geolocated

Location is refined & actionable items are identified

Original: “Fanm gen tranche pou fè yon pitit nan Delmas 31”

Translation: Woman in labor having a child at Delmas 31

Location: 18.495746829274168, 72.31849193572998

Category: Emergency

Frequency of text messages to U-Report in Nigeria

[Long, Munro, Gaylord, Asfaw & Wheeler. (2015). Classification Models for New Language Communities: Building Domain-Specific Message Categorization. *ACM-DEV*]

DeepLearning.AI

AI and Public Health

Explore Phase

Problem Definition and Stakeholders

1,000 days project

[Wikimedia Commons, 2006. Nigeria location]

Nutrition and Healthcare for mothers

Pregnancy

2 years old

AI4G Framework: Explore Phase

AI4G Framework: Explore Phase

Deploy

Explore

1. Engage stakeholders
2. Define the problem
3. Determine if AI could add value

A good problem statement should:

1. Define the problem you are hoping to address.
2. Identify key stakeholders.
3. Give an idea of what success looks like.
4. Not mention the specific technology you aim to deploy.

Explore: Define the problem

"It's hard to process large amounts of text messages in different languages."

Explore | Design | Implement | Evaluate

1. Engage stakeholders
2. Define the problem
3. Determine if AI could add value

Does not identify key stakeholders

Explore: Define the problem

"It's hard to process large amounts of text messages in different languages."

Explore

1. Engage stakeholders
2. Define the problem
3. Determine if AI could add value

"**Healthcare providers need to communicate directly with mothers in the community** to monitor their health and the health of their babies, and they are faced with the challenge of processing large volumes of text messages in multiple languages."

Identifies key stakeholders

Does not give an idea of what success looks like

Explore: Define the problem

"It's hard to process large amounts of text messages in different languages."

Explore

1. Engage stakeholders
2. Define the problem
3. Determine if AI could add value

"**Healthcare providers need to communicate directly with mothers in the community** to monitor their health and the health of their babies, and they are faced with the challenge of processing large volumes of text messages in multiple languages."

"Healthcare providers need to communicate directly with mothers in the community via surveys to monitor their health and the health of their babies. To do this, they need to be able to **quickly process a large volume of incoming text messages in multiple languages** including survey responses and other unrelated messages from the community."

Explore: Define the problem

Explore

1. Engage stakeholders
2. Define the problem
3. Determine if AI could add value

A good problem statement should:

1. Be clear, concise, and specific.
2. Identify key stakeholders.
3. Give an idea of what success looks like.
4. Not mention the specific technology you aim to deploy.

"Healthcare providers need to communicate directly with mothers in the community via surveys to monitor their health and the health of their babies. To do this, they need to be able to **quickly process a large volume of incoming text messages in multiple languages** including survey responses and other unrelated messages from the community."

DeepLearning.AI

AI and Public Health

Explore Phase

Could AI Add Value?

Supervised learning summary

Input A	Output B	Application
Road image	Cyclist / not cyclist	Image recognition

Cyclist

Not cyclist

Supervised learning summary

Input A	Output B	Application
Road image	Cyclist / not cyclist	Image recognition
Cry recordings	Healthy / not healthy	Speech recognition

Healthy

Not healthy

AI for Good framework

1. Engage stakeholders
2. Define the problem
3. Determine if AI could add value

Explore: Does AI add value?

1. Engage stakeholders
2. Define the problem
3. Determine if AI could add value

Do no harm

1. Message contain personally identifiable information that should be kept **private and secure**
2. Ensure that responses are accurate.

Explore phase checkpoint

- What is the specific problem you are addressing?
- Who are the stakeholders?
- Do you have access to or can you collect the necessary data?
- Could AI add value? Where and how specifically?
- How does the “do no harm” principle come into play?

Explore phase checkpoint

"Healthcare providers need to communicate directly with mothers in the community via surveys to monitor their health and the health of their babies. To do this, they need to be able to **quickly process a large volume of incoming text messages in multiple languages** including survey responses and other unrelated messages from the community."

- What is the specific problem you are addressing? ✓
- Who are the stakeholders?
- Do you have access to or can you collect the necessary data?
- Could AI add value? Where and how specifically?
- How does the "do no harm" principle come into play?

Explore phase checkpoint

- What is the specific problem you are addressing? ✓
- Who are the stakeholders? ✓
- Do you have access to or can you collect the necessary data?
- Could AI add value? Where and how specifically?
- How does the “do no harm” principle come into play?

Explore phase checkpoint

Checkpoint

- What is the specific problem you are addressing? ✓
- Who are the stakeholders? ✓
- Do you have access to or can you collect the necessary data? ✓
- Could AI add value? Where and how specifically?
- How does the "do no harm" principle come into play?

Explore phase checkpoint

Checkpoint

- What is the specific problem you are addressing? ✓
- Who are the stakeholders? ✓
- Do you have access to or can you collect the necessary data? ✓
- Could AI add value? Where and how specifically? ✓
- How does the "do no harm" principle come into play?

Explore phase checkpoint

- What is the specific problem you are addressing? ✓
- Who are the stakeholders? ✓
- Do you have access to or can you collect the necessary data? ✓
- Could AI add value? Where and how specifically? ✓
- How does the “do no harm” principle come into play? ✓

DeepLearning.AI

AI and Public Health

Design Phase

AI for Good framework

1. Prototype your solution
2. Ensure data privacy
3. Design the user experience

AI for Good framework

1. Prototype your solution
2. Ensure data privacy
3. Design the user experience

AI for Good framework

1. Prototype your solution
2. Ensure data privacy
3. Design the user experience

Design phase

Task: Determine the minimum number of labels/categories

Explore | **Design** | Implement | Evaluate

1. Prototype your solution
2. Ensure data privacy
3. Design the user experience

Input A → **Output B**

Text Messages

Categorized Messages

Text message classification

Input → **Output**

Text Messages

Administrative
related to
the clinic

1. Patient-related
2. Clinic-admin
3. Technological

Requests
from Community
Health Workers

1. Response
2. Request for doctor
3. Medical advice

Illness
monitoring bodies

1. TB
2. HIV
3. Death

State-of-the-art Model

- ⊕ Predictions are more accurate
- ⊖ May take longer to make predictions
- ⊖ May take longer to retrain the model

Simpler Model

- More reliable
- Easier to understand and interpret
- Able to retrain the model faster

Does AI really add value?

Design phase

1. Prototype your solution
2. Ensure data privacy
3. Design the user experience

Medical information contained PII

- Access to messages was limited
- Messages were not downloaded or stored

Design phase

1. Prototype your solution
2. Ensure data privacy
3. Design the user experience

Goals:

- Process more messages
- Reduce response time

Design phase checkpoint

- How will you address issues with imbalances, biases, privacy, or other concerns with your data?
- What kind of model will you implement, and how will you measure its performance?
- How will your design address the problem you set out to work on?
- How will the end user interact with your system?

Data considerations: Privacy & security

Explore | **Design** | Implement | Evaluate

Only essential medical staff had access to the data

Text message classification

Input → **Output**

Text Messages

Administrative
related to
the clinic

1. Patient-related
2. Clinic-admin
3. Technological

Requests
from Community
Health Workers

1. Response
2. Request for doctor
3. Medical advice

Illness
monitoring bodies

1. TB
2. HIV
3. Death

DeepLearning.AI

AI and Public Health

Implement Phase

AI for Good framework

1. Productionize AI models
2. Integrate the user experience
3. Test with end users

AI for Good framework

1. Productionize AI models
2. Integrate the user experience
3. Test with end users

AI for Good framework

1. Productionize AI models
2. Integrate the user experience
3. Test with end users

Maternal & infant health classification

Implement phase checkpoint

- Is your model performance acceptable?
- Are end users able to successfully use your system?

AI and Public Health

DeepLearning.AI

Evaluate Phase

AI for Good framework

1. Measure project impact
2. Communicate results
3. Determine next steps

AI for Good framework

1. Measure project impact
2. Communicate results
3. Determine next steps

AI for Good framework

1. Measure project impact
2. Communicate results
3. Determine next steps

AI for Good framework

"**Healthcare providers** need to communicate directly with mothers in the community via surveys to monitor their health and the health of their babies. To do this, they need to be able to **quickly process a large volume of incoming text messages in multiple languages** including survey responses and other unrelated messages from the community."

1. Measure project impact
2. Communicate results
3. Determine next steps

AI is only one part of any potential solution

Maternal & infant health classification

AI for Good framework

Bad user experience:
Too time consuming for clinic staff

Manual annotations discontinued

Conclusion: Project was discontinued

1. Measure project impact
2. Communicate results
3. Determine next steps

U Report VOICE MATTERS

TRANSLATORS WITHOUT BORDERS

“Do no harm”

Everyone impacted by a project is left improved or unchanged.

AI for Good framework

1. Measure project impact
2. Communicate results
3. Determine next steps

As of 2023, U-Report is active in 68 countries, serving over 11 million users worldwide

DeepLearning.AI

AI and Public Health

Iva Gumnishka – Building an Ethical Supply Chain

ONLY

Jacket
Fleece
Orange
Jacket
Hoodie
Red
Pants
Jeans
Blue
Pants
Cloth
Brown
Pants
Jeans
Black

Trash and litter analysis

Robotic surgery

Damage detection

Need human input!

Our role

humans collect, generate and annotate the training data

1st stage:
model training

2nd stage:
model testing

humans give feedback for reinforcement learning

humans review the errors and provide insights as an AI audit

4th stage:
model auditing

3rd stage:
model deployment

humans monitor the AI model and edge cases in real time

DeepLearning.AI
UKRAINE
BULGARIA
TURKEY
PORTUGAL
LEBANON
SYRIA
IRAQ
AFGHANISTAN
YEMEN
COLOMBIA
DRC

2. Are you using a **diverse group of humans** to provide this input?

W2 Lesson 2

Exploring Air Quality

DeepLearning.AI

AI and Public Health

Air Quality

Estimating pollution levels in Bogotá

AI for Good framework

Explore → Design → Implement → Evaluate

Estimating pollution levels in Bogotá

AI for Good framework

Explore → Design → Implement → Evaluate

Estimating pollution levels in Bogotá

AI for Good framework

Explore | **Design** | Implement | Evaluate

Estimating pollution levels in Bogotá

AI for Good framework

Explore → Design → **Implement** → Evaluate

Estimating pollution levels in Bogotá

AI for Good framework

Explore → Design → Implement → Evaluate

Estimating pollution levels in Bogotá

[Wikimedia Commons, 2009. Colombia location.]

DeepLearning.AI

AI and Public Health

Sources of Air Pollution

Air quality and public health

Air pollution

Outdoor

Modern ambient air pollution

Death rate due to air pollution is 50% higher today than 20 years ago

Most common air pollutants

- Particulate matter (PM)
- Ozone
- Lead
- Oxides of nitrogen, sulfur, carbon

Most common air pollutants

- **Particulate matter ($PM_{2.5}$)**
- Ozone
- Lead
- Oxides of nitrogen, sulfur, carbon

Particulate matter 2.5 ($PM_{2.5}$)

Human hair
50-70μm
(microns)

PM 2.5
2.5μm
(microns)

20-30 PM2.5 particles span the width of one human hair!

- Chronic lung problems
- Lung cancer
- Cardiovascular diseases

DeepLearning.AI

AI and Public Health

Measuring Air Quality

DeepLearning.AI

AI and Public Health

Bogotá Air Quality Monitoring Network (RMCAB)

Bogotá, Colombia

[Wikimedia Commons, 2009. Colombia location.]

Bogotá Air Quality Monitoring Network

Measures Pollutants

- Particulate Matter
- Ozone
- Sulfur Oxide
- Nitrogen Oxide
- Carbon Oxide
- Atmospheric Conditions

Air Quality Sensors

RMCAB in Bogotá, Colombia

PurpleAir Sensors in California, U.S.A

The U.N. sustainable development goal targets

2030 Goal: 70% of the city within the recommended healthy limits for air quality.

Missing data

Under maintenance

No data available for these sensors

Partial data

No data in-between the sensors

DeepLearning.AI

AI and Public Health

Air Quality

Explore Phase

AI for Good framework

1. Engage stakeholders
2. Define the problem
3. Determine if AI could add value

AI for Good framework

1. Engage stakeholders
2. Define the problem
3. Determine if AI could add value

AI for Good framework

Deploy

Explore

Desi

lem

aluate

1. Engage stakeholders
2. Define the problem
3. Determine if AI could add value

No data available for these sensors

No data in-between the sensors

AI for Good framework

Deploy

Explore

1. Engage stakeholders
2. Define the problem
3. Determine if AI could add value

"Public health professionals working with the city of Bogotá need to be able to provide **real time estimates of air quality** throughout the city so that citizens can be aware of any health risks due to poor air quality and plan their outdoor activities accordingly."

DeepLearning.AI

AI and Public Health

Air Quality

Explore and Visualize Missing Data

5. Make box plots of pollutants across all sensor stations

Box plots are another way of investigating the distribution of your data. With a boxplot, your range of data values runs along the vertical axis and each "box with whiskers" provides you with information on where the median (middle of the box) and middle 50% (vertical range of the box) of your data lies. Individual data points outside the middle 50% are also plotted on the vertical axis. In this case, there's one box plot each for every sensor station.

Run the next cell to generate box plots and use the pulldown menu to look at the distributions for different pollutants.

Horizontal line shows the median of the distribution

In [5]:
```
# Generate boxplots of pollutant values for each sensor station
utils.create_boxplot(raw_data)
```

Pollutant PM2.5

6. Investigate scatter plots of different pollutants and look for correlation

Scatterplots are useful for looking at the relationship (or lack thereof) between two variables. Run the cell below to generate scatterplots of one pollutant plotted against another. Use the pulldown menus to investigate the relationship between different pollutants. Do any of the pollutants appear to be correlated with one another? The colors on the plot show the concentration of the points. the closer towards the red, the more points are concentrated in the given spot.

5. Make box plots of pollutants across all sensor stations

Box plots are another way of investigating the distribution of your data. With a boxplot, your range of data values runs along the vertical axis and each "box with whiskers" provides you with information on where the median (middle of the box) and middle 50% (vertical range of the box) of your data lies. Individual data points outside the middle 50% are also plotted on the vertical axis. In this case, there's one box plot each for every sensor stati

Run the next cell to generate box plots and use the pulldown menu to look at the distributions for different pollutants.

Shaded box shows 50% of the data around the mean

In [5]:
```
# Generate boxplots of pollutant values for each sensor station
utils.create_boxplot(raw_data)
```

Pollutant PM2.5

6. Investigate scatter plots of different pollutants and look for correlation

Scatterplots are useful for looking at the relationship (or lack thereof) between two variables. Run the cell below to generate scatterplots of one pollutant plotted against another. Use the pulldown menus to investigate the relationship between different pollutants. Do any of the pollutants appear to be correlated with one another? The colors on the plot show the concentration of the points. the closer towards the red, the more points are concentrated in the given spot.

5. Make box plots of pollutants across all sensor stations

Box plots are another way of investigating the distribution of your data. With a boxplot, your range of data values runs along the vertical axis and each "box with whiskers" provides you with information on where the median (middle of the box) and middle 50% (vertical range of the box) of your data lies. Individual data points outside the middle 50% are also plotted on the vertical axis. In this case, there's one box plot each for every sensor station.

Run the next cell to generate box plots and use the pulldown menu to look at the distributions for different pollutants.

In [5]:
```
# Generate boxplots of pollutant values for each sensor station
utils.create_boxplot(raw_data)
```

Pollutant PM2.5

Third Quartile (Q3)
First Quartile (Q1)

6. Investigate scatter plots of different pollutants and look for correlation

Scatterplots are useful for looking at the relationship (or lack thereof) between two variables. Run the cell below to generate scatterplots of one pollutant plotted against another. Use the pulldown menus to investigate the relationship between different pollutants. Do any of the pollutants appear to be correlated with one another? The colors on the plot show the concentration of the points. the closer towards the red, the more points are concentrated in the given spot.

5. Make box plots of pollutants across all sensor stations

Box plots are another way of investigating the distribution of your data. With a boxplot, your range of data values runs along the vertical axis and each "box with whiskers" provides you with information on where the median (middle of the box) and middle 50% (vertical range of the box) of your data lies. Individual data points outside the middle 50% are also plotted on the vertical axis. In this case, there's one box plot each for every sen

Run the next cell to generate box plots and use the pulldown menu to look at the distributions for different pollutants.

The data is sparse at high values of PM2.5 concentration

In [5]:
```
# Generate boxplots of pollutant values for each sensor station
utils.create_boxplot(raw_data)
```

Pollutant PM2.5

6. Investigate scatter plots of different pollutants and look for correlation

Scatterplots are useful for looking at the relationship (or lack thereof) between two variables. Run the cell below to generate scatterplots of one pollutant plotted against another. Use the pulldown menus to investigate the relationship between different pollutants. Do any of the pollutants appear to be correlated with one another? The colors on the plot show the concentration of the points. the closer towards the red, the more points are concentrated in the given spot.

DeepLearning.AI

AI and Public Health

Air Quality

Explore Correlations

Run Markdown

Run Markdown

DeepLearning.AI

AI and Public Health

Air Quality

Explore Phase Checkpoint

Explore phase checkpoint

- What is the specific problem you are addressing?
- Who are the stakeholders?
- Do you have access to or can you collect the necessary data?
- Could AI add value? Where and how specifically?
- How does the “do no harm” principle come into play?

Explore phase checkpoint

Checkpoint

- What is the specific problem you are addressing?

"Public health professionals working with the city of Bogotá need to be able to provide **real time estimates of air quality** throughout the city so that citizens can be aware of any health risks due to poor air quality and plan their outdoor activities accordingly."

Explore phase checkpoint

Checkpoint

- What is the specific problem you are addressing? ✓
- Who are the stakeholders? ✓

Explore phase checkpoint

	PM10	PM2.5	NO	NO2	NOX	CO	OZONE	Station	DateTime
0	56.6	32.7	7.504	15.962	23.493	0.44924	2.431	USM	2021-01-01 00:00:00
1	59.3	39.3	16.560	17.866	34.426	0.69832	1.121	USM	2021-01-01 01:00:00
2	96.4	70.8	22.989	17.802	40.791	0.88243	1.172	USM	2021-01-01 02:00:00
3	108.3	81.0	3.704	9.886	13.591	0.29549	6.565	USM	2021-01-01 03:00:00
4	87.7	56.1	2.098	9.272	11.371	0.16621	9.513	USM	2021-01-01 04:00:00

- What is the specific problem you are addressing? ✓
- Who are the stakeholders? ✓
- Do you have access to or can you collect the necessary data? ✓
- Could AI add value? Where and how specifically? ✓

Explore phase checkpoint

Explore → Design → Implem

Checkpoint

- What is the specific problem you are addressing? ✓
- Who are the stakeholders? ✓
- Do you have access to or can you collect the necessary data? ✓
- Could AI add value? Where and how specifically?

Explore phase checkpoint

Checkpoint

- What is the specific problem you are addressing? ✓
- Who are the stakeholders? ✓
- Do you have access to or can you collect the necessary data? ✓
- Could AI add value? Where and how specifically? ✓

Explore phase checkpoint

- What is the specific problem you are addressing? ✓
- Who are the stakeholders? ✓
- Do you have access to or can you collect the necessary data? ✓
- Could AI add value? Where and how specifically? ✓
- How does the “do no harm” principle come into play? ✓