

Natural-Language Processing (NLP)

Kuliah-1

9/20/2022

Zulkaida Akbar

Today's goal

- A brief introduction to NLP
- Jupyter Notebook tutorial
- Speech analysis in Bahasa Indonesia
- Speech analysis in English
- Vader Sentiment Analysis (If time permit)
- Homework

Natural Language

English language

Chinese language

~~Python language~~

Processing

How a computer
carries out instructions

How to deal with text data

Artificial Intelligence

A computer performing tasks that a human can do



Natural Language Processing



Text: "I love your hats so much! ..."

Length: 1 minute

Rep: Sally



Text: "My shirt doesn't fit. ..."

Length: 5 minutes

Rep: Sally



Text: "What are your newest hats? ..."

Length: 3 minutes

Rep: Bobby



Example-1

Sentiment Analysis

Example-2

Topic Modeling



Here's the agenda for our team meeting on Friday.

project

The final cost of the project comes to \$1MM.

project \$\$\$

I've attached the receipt of my travel expenses.

\$\$\$

Can you forward that request to my admin?

project

Example-3

Text Generation



You are amazing.

Take the road less traveled.

You do you.

Do what makes you happy.

What's hard makes you stronger.

Love is a journey.

NLP Examples



Sentiment Analysis

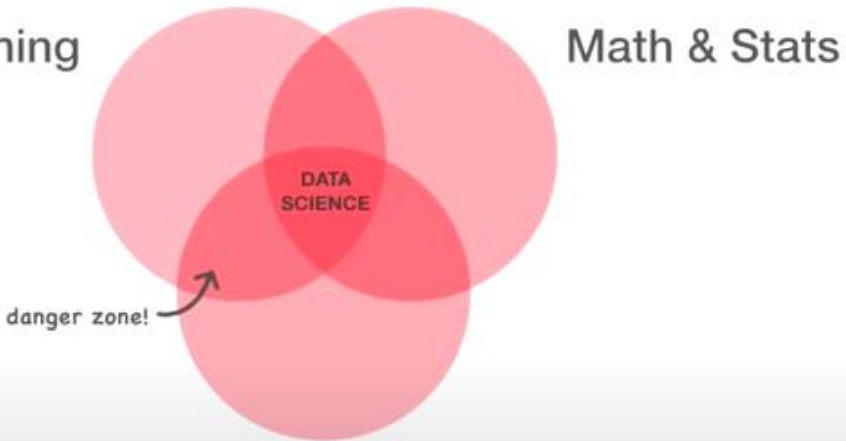


Topic Modeling



Text Generation

Data Science Venn Diagram



NLP in Python

Programming	<i>Data NLP</i>	pandas, sklearn, re nltk, TextBlob, gensim
Math & Stats	<i>Clean EDA* NLP</i>	corpus, document-term matrix word counts sentiment analysis, topic modeling, text generation
Communication	<i>Design Domain</i>	scope, visualize, extract insights expertise

Data Science Workflow

1. Start with a Question
2. Get & Clean the Data
3. Perform EDA
4. Apply Techniques
5. Share Insights

A Simple Example

1. Start with a Question

2. Get & Clean the Data

3. Perform EDA

4. Apply Techniques

5. Share Insights

"If I study more, will I get a higher grade?"

1. Start with a Question

2. Get & Clean the Data

3. Perform EDA

4. Apply Techniques

5. Share Insights

"If I study more, will I get a higher grade?"

Student	Hours Studied	Grade
Alice	20	90
Bob	5	70
Charlie	10	96
David	15	82
Eve	two	62
Frank	16	87
Grace	22	998

Need cleaning

A Simple Example

1. Start with a Question

2. Get & Clean the Data

3. Perform EDA

4. Apply Techniques

5. Share Insights

"If I study more, will I get a higher grade?"



1. Start with a Question

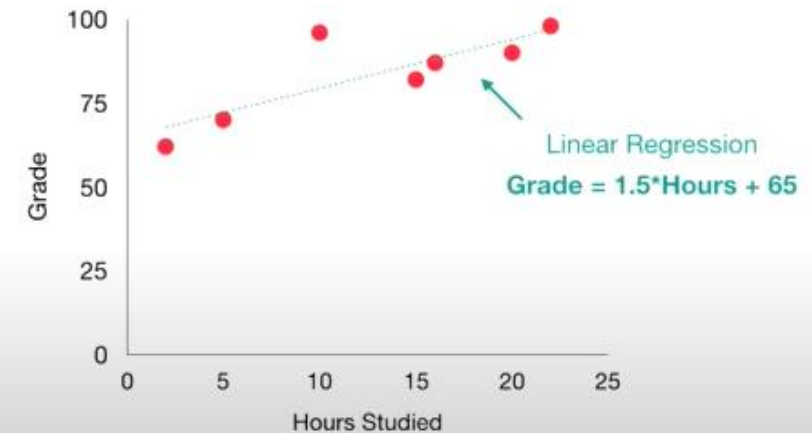
2. Get & Clean the Data

3. Perform EDA

4. Apply Techniques

5. Share Insights

"If I study more, will I get a higher grade?"



1. Start with a Question

2. Get & Clean the Data

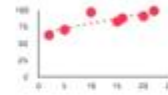
3. Perform EDA

4. Apply Techniques

5. Share Insights

"If I study more, will I get a higher grade?"

Yes, there is a **positive correlation** between the number of hours you study and the grade you will get.

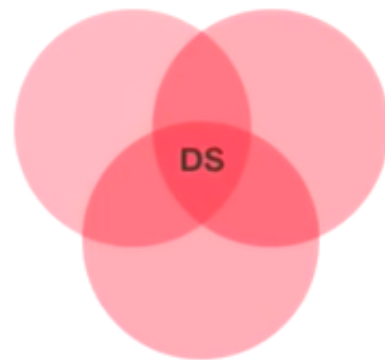


Specifically, the relationship is: **Grade = 1.5 * Hours + 65**
So if you study **10 hours**, you can expect to get an **80**.

However, Charlie is a smarty pants and is inflating the grade estimate. You'll probably get **slightly less than 80**.

Summary

Analytics:
Using data to
make
decisions



1. Start with a Question

2. Get & Clean the Data

3. Perform EDA

4. Apply Techniques

5. Share Insights

Purpose

Venn Diagram

Workflow

Jupyter Notebook Tutorial:

- Download anaconda: <https://www.anaconda.com/products/distribution>
- Launch Jupyter Notebook

Windows: Open the Anaconda Navigator program. You should see the Jupyter Notebook logo. Below the logo, click Launch. A browser window should open up. In the browser window, navigate to the location of the saved Jupyter Notebook files and open 0-Hello-World.ipynb. Follow the instructions in the notebook.

Mac/Linux: Open a terminal. Type `jupyter notebook`. A browser should open up. In the browser window, navigate to the location of the saved Jupyter Notebook files and open 0-Hello-World.ipynb. Follow the instructions in the notebook.

- Download a few more packages

There are a few additional packages we'll be using during the tutorial that are not included when you download Anaconda - wordcloud, textblob and gensim.

Windows: Open the Anaconda Prompt program. You should see a black window pop up. Type `conda install -c conda-forge wordcloud` to download wordcloud. You will be asked whether you want to proceed or not. Type `y` for yes. Once that is done, type `conda install -c conda-forge textblob` to download textblob and `y` to proceed, and type `conda install -c conda-forge gensim` to download gensim and `y` to proceed.

Mac/Linux: Your terminal should already be open. Type `command-t` to open a new tab. Type `conda install -c conda-forge wordcloud` to download wordcloud. You will be asked whether you want to proceed or not. Type `y` for yes. Once that is done, type `conda install -c conda-forge textblob` to download textblob and `y` to proceed, and type `conda install -c conda-forge gensim` to download gensim and `y` to proceed.

If you have any issues, please email me at adashofdata@gmail.com or come talk to me before the start of the tutorial.

- Open: <https://www.dataquest.io/blog/jupyter-notebook-tutorial/>
- Enjoy the tutorial

Speech analysis and sentiment analysis Tutorial:

- Download github: <https://github.com/zulkaidaakbar/NLP-STTCipasung2022/tree/main/20-September-2022>
- Follow the instruction
- Enjoy the tutorial