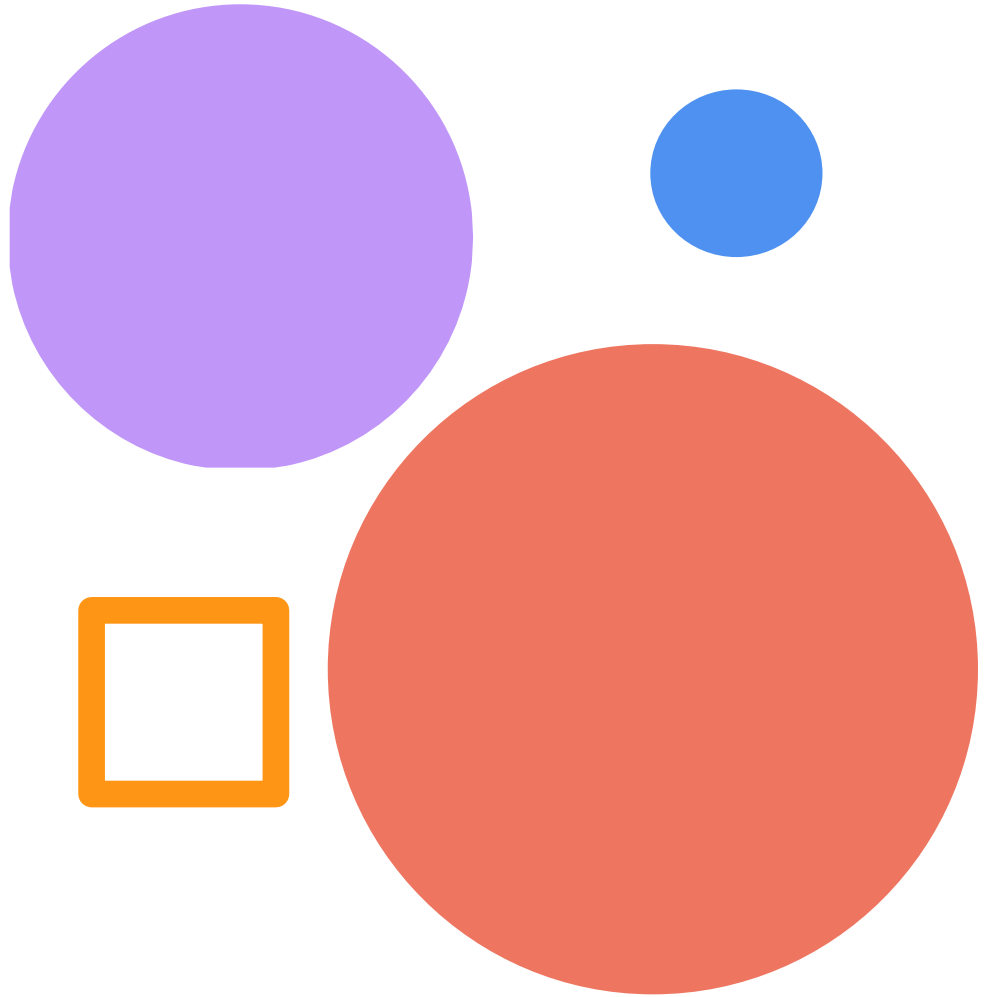
Automated News Categorizing

Proposal Title Presentation

Content

- Overview
- Problem Statement
- Objective of Research
- Scope
- Significance of Study
- Software and Hardware Are Requirement
- Relate Work
- Methodology

Overview

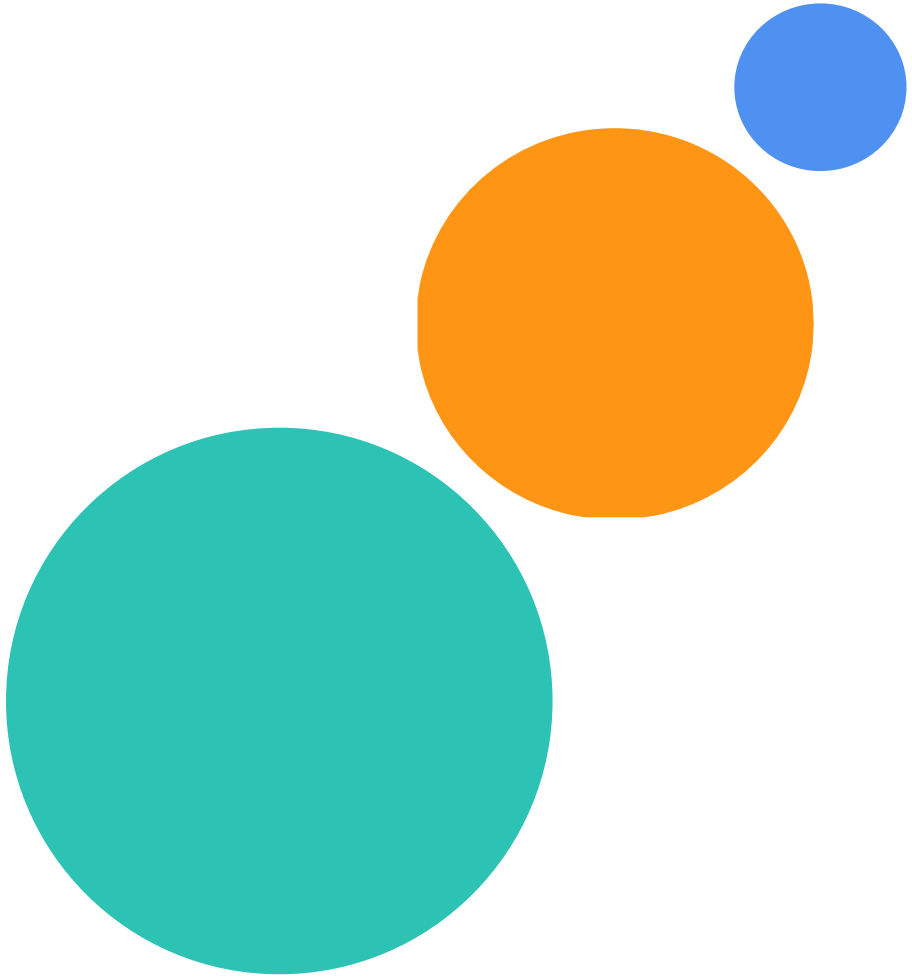


Problem Statement

Objective of Research

- To use the machine learning to categorize news and deployment by using technique of data visualization and data storytelling.
- To visualize the categorized news into news timeline visualization in form of data storytelling on a web application.

Scope



Significance of Study

Developer

- To gets technique skill of data analysis in term of using machine learning to apply with news categorizing.
- To deployment the result on web application by using python.

User

- User can follow the trend of news are viral in social during that time.
- User gets quick access to relevant news and topics of interest.

Software and Hardware are Requirement

Software Requirement

Resource

- News from Bangkok post agency website.

Data Mining Process

- Python
- Jupyter Notebook
- Data Mining Libraries
  - Requests
  - Beautiful Soup 4
  - Pandas
  - NumPy
  - re (Regular Expression)

Software Requirement

Data Visualization Process

- Web Application Language
  - HTML
  - JavaScript
  - Css
  - Python
- Data Visualization Libraries
  - Django
  - Matplotlib
  - Seaborn

Data Base

- MySQL
- Data Integration Format
  - JSON

Hardware Requirement

Personal Computer

- Asus VivoBook 15 x512da
- HP Pavilion Power 15-cb035TX

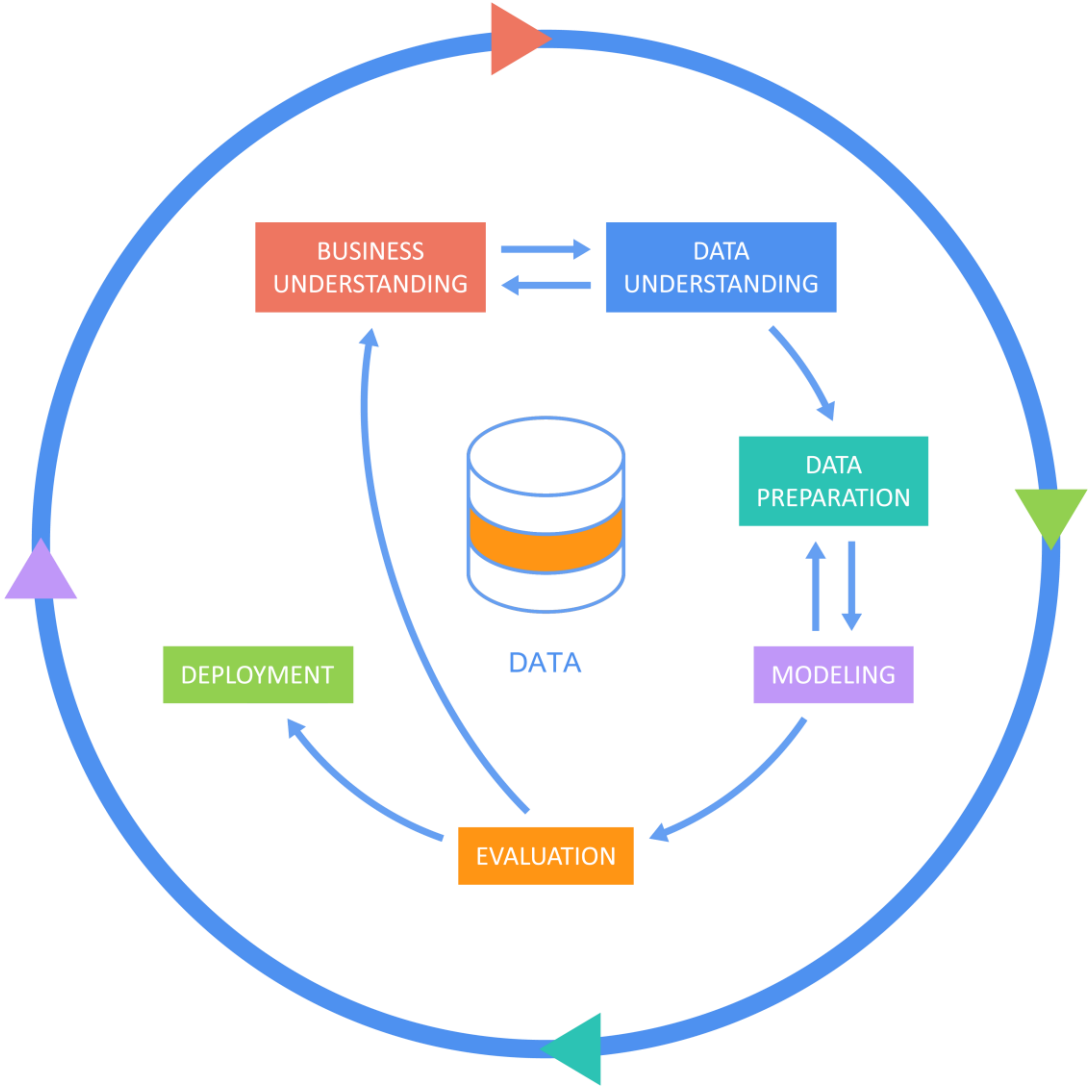
Relate Work

- Automated Semantic Categorization of News Headlines Using Ensemble Machine Learning: A Comparative Study. (Bogery et al., 2019)

- Automated Text Tagging of Arabic News Articles Using Ensemble Deep Learning Models. (Elnagar et al., 2019)

Methodology

CRISP-DM


BUSINESS UNDERSTANDING
DATA UNDERSTANDING
DATA PREPARATION
MODELING
EVALUATION
DEPLOYMENT
DATA

Methodology

Business Understanding

Categorizing news automatically by using machine learning SVM (Support Vector Machine).

Data Understanding

- Resource
  - Bangkok post agency website.
- File extension
  - JSON.

Data Preparation

NLP (Natural Language Processing).

Methodology

Modelling

SVM (Support Vector Machine).

Evaluation

Model Evaluation.

Deployment

Web Application.

Thank You :)