Contents

[**1.** **Topic Modelling on online news** 2](#_Toc83000790)

[**Topic 1: Pandemic on Political** 5](#_Toc83000791)

[**Topic 2: Item Price Issue** 6](#_Toc83000792)

[**Topic 3: Operation/Work Issue** 7](#_Toc83000793)

[**Topic 4: Oil-related issue/ Subsidised** 8](#_Toc83000794)

[**Topic 5: Vaccination News** 9](#_Toc83000795)

[**2.** **Topic Modelling on social media** 10](#_Toc83000796)

[**Topic 1: Harga** 12](#_Toc83000797)

[**Topic 2: Negative comments** 13](#_Toc83000798)

[**Topic 3: Penguatkuasaan** 14](#_Toc83000799)

[**Topic 4: Delete Tweet** 15](#_Toc83000800)

## **Topic Modelling on online news**

Figure 1: Methodology in topic modeling for e-learning

Chart, bar chart

Description automatically generated

Figure 2: News distribution by months (2021)

1. Total news collected from 1 January 2021 until 18 September 2021 is 5584. Figure 2 shows the distribution of the news collected.
2. After preprocessing, only 5259 news is selected for analysis.
3. Non-Negative Matrix Factorization (NMF) Model is used to find the hidden topic on the documents/ each news
4. 5000 words as features with 5 target topics.
5. The result of topic modeling is shown in Figure 4.

Bar chart

Description automatically generated with low confidence

Figure 3: The highly correlated word given by the topic

Chart, bar chart

Description automatically generated

Figure 4: Distribution of topics across the news data

### **Topic 1: Pandemic on Political**

Text

Description automatically generated with medium confidence

Word cloud on Topic 1

Graphical user interface

Description automatically generated with medium confidence Graphical user interface, table

Description automatically generated

Bigrams word frequency on Topic 1

### **Topic 2: Item Price Issue**

Text

Description automatically generated

Word cloud on Topic 2

Table

Description automatically generated with low confidence Graphical user interface, table

Description automatically generated with medium confidence

Bigrams word frequency on Topic 2

Note: 019-279 4317 is the KPDNHEP hotline

### **Topic 3: Operation/Work Issue**

Text

Description automatically generated

Word cloud on Topic 3

Graphical user interface, application

Description automatically generated Graphical user interface, table

Description automatically generated

Bigrams word frequency on Topic 3

### **Topic 4: Oil-related issue/ Subsidised**

Text

Description automatically generated

Word cloud on Topic 4

Graphical user interface, table

Description automatically generated Table

Description automatically generated

Bigrams word frequency on Topic 4

### **Topic 5: Vaccination News**

Text

Description automatically generated

Word cloud on Topic 5

Table

Description automatically generated Table

Description automatically generated

Bigrams word frequency on Topic 5

## **Topic Modelling on social media**

1. Twitter tweets are collected from 1 January 2021 until 18 September 2021. The total tweets collected are 1200.
2. Only tweets that mentioned account that is related to KPDNHEP are collected. A total of 14 accounts that are related to KPDNHEP. The accounts are listed as follows:
3. @kpdnhep
4. @KPDNHEPKelantan
5. @KPDNHEPPerak
6. @kpdnhepjohor
7. @kpdnheptrg
8. @KpdnhepPP
9. @Nforceteam1
10. @KpdnhepM
11. @negeri\_mykpdnkk
12. @kppmkpdnkk
13. @penang\_mykpdnkk
14. @KpdnhepK
15. @kpdnhepsabah
16. @kps\_mykpdnhep

1. After preprocessing, only 1002 tweets are used for topic modeling analysis. On a first look, most tweets mentioned the official Twitter page which is @kpdnhep. The distribution of the tweets are as follows:

Text

Description automatically generated

Figure 5: Methodology in topic modeling for e-learning

Chart, histogram

Description automatically generated

Figure 6: Distribution of topics across the news data

### **Topic 1: Harga**

A picture containing text

Description automatically generated

Word cloud on Topic 1

Graphical user interface

Description automatically generated with low confidence

Bigrams word frequency on Topic 1

### **Topic 2: Negative comments**

A picture containing text, newspaper

Description automatically generated

Word cloud on Topic 2

Table

Description automatically generated

Bigrams word frequency on Topic 2

### **Topic 3: Penguatkuasaan**

Text

Description automatically generated

Word cloud on Topic 3

Table

Description automatically generated

Bigrams word frequency on Topic 3

### **Topic 4: Delete Tweet**

Negative comments on a deleted tweet by KPDNHEP. KPDNHEP’s tweet triggers some Twitter users.

Reference: <https://www.freemalaysiatoday.com/category/bahasa/tempatan/2021/09/14/itu-advokasi-pengguna-bukan-kempen-berat-sebelah-kata-kpdnhep/>

Text

Description automatically generated with medium confidence

Word cloud on Topic 4

Graphical user interface

Description automatically generated with low confidence

Bigrams word frequency on Topic 4