

## APPENDIX

### INDEX

```
<!DOCTYPE html>
<html>
<head>
  <title>News Clustering</title>
  <!--<link rel="stylesheet" href="css/jquery-ui.min.css">-->
  <link rel="stylesheet"
href="//code.jquery.com/ui/1.12.1/themes/base/jquery-ui.css">
  <style type="text/css">
    div{
      padding: 5px;
    }
    .container{
      display: flex;
    }
    .fixed{
      border: 2px solid black;
      width: 250px;
    }
    .flex-item{
      flex-grow: 1;
    }
  </style>
</head>
<body>
  <div class="container">
    <div class="fixed">
      <form role="form" id="main" method="post"
enctype="multipart/form-data">
        Pilih artikel yang akan dicari
kelompoknya<br />
        <input type="file" name="files[]" required
multiple>
        <br /><br />
        Pilih situs berita<br />
        <select name="select[]" multiple>
          <option value="eko">Kompas
Ekonomi</option>
          <option value="oto">Kompas
Otomotif</option>
          <option value="tekno">Kompas
Tekno</option>
          <option value="travel">Kompas
Travel</option>
        </select>
        <br /><br />
        Pilih waktu berita<br />
        Dari<br />
        <input type="text" class="datepicker"
name="startdate">
```

```

                <br /><br />
                Sampai<br          /><input          type="text"
class="datepicker" name="enddate">
                <br /><br />
                <input          type="submit"          name="action"
value="Submit">
                </form>
            </div>
            <div class="flex-item">
                <div id="data">
            </div>
        </div>
        <script src="js/jquery.min.js"></script>
        <script src="js/jquery-ui.min.js"></script>
        <script>
        $( function() {
        $( ".datepicker" ).datepicker();
        } );
        //form Submit action
        $("#main").submit(function(evt){
            evt.preventDefault();
            var formData = new FormData($(this)[0]);
            $.ajax({
                url: 'process.php',
                type: 'POST',
                data: formData,
                async: false,
                cache: false,
                contentType: false,
                enctype: 'multipart/form-data',
                processData: false,
                success: function (response) {
                    document.getElementById("data").innerHTML=response;
                }
            });
            return false;
        });
        </script>
    </body>
</html>

```

#### INCLUDE FROM CLASSES

```

include('class/simplehtmlDom/simple_html_dom.php');
include('class/cURL.php');
include('class/webscraping.php');
include('class/tokenisasi.php');
include('class/stopword.php');
include('class/stemmer.php');
include('class/termweight.php');
include('class/cluster.php');

```

#### READ INPUTED NEWS ARTICLE FILE

```

$filex = $_FILES['files'];

for($cc=0;$cc<count($filex['name']);$cc++)
{
    $path = 'datac/';
    $filename = $filex['name'][$cc];
    $target = $path.$filename;

    if($filex['error'][$cc])
    {
        echo "Error : ".$filex['error'][$cc]."<br />";
    }
    else
    {
        move_uploaded_file($filex['tmp_name'][$cc], $target);

        $k[$cc] = file_get_contents($target);
    }
}

```

**GET NEWS FROM KOMPAS.COM FROM SPECIFIC DATE AND CATEGORY**

```

$startdate = $_POST['startdate'];
$enddate = $_POST['enddate'];

$begin = new DateTime($startdate);
$end = new DateTime($enddate);
$end = $end->modify( '+1 day' );

$interval = new DateInterval('P1D');
$daterange = new DatePeriod($begin, $interval , $end);

foreach($daterange as $date)
{
    $tgl = $date->format("d");
    $bln = $date->format("m");
    $thn = $date->format("Y");
    //echo "Artikel tanggal : $tgl-$bln-$thn."<br>";

    if(in_array('eko', $_POST['select']))
    {
        $scrap = new Webscraping();
        /*$deko = $scrap-
>getDetikArticles("https://finance.detik.com/indeks?date=$bln
%2F$stgl%2F$sthn"); //detik finance
        for ($i=0; $i <count($deko) ; $i++) {
            $full[] = $deko[$i];
            $sarr = explode("\n", $deko[$i]);
            $judul[] = $sarr[0];
            $isi[] = $sarr[1];
        }*/
    }
}

```

```

        $keko = $scrap-
>getKompasArticles("http://bisniskeuangan.kompas.com/search/$thn-
$b1n-$t1"); //kompas bisnis
    for ($i=0; $i <count($keko) ; $i++) {
        $full[] = $keko[$i];
        $arr = explode("\n", $keko[$i]);
        $judul[] = $arr[0];
        $isi[] = $arr[1];
    }
}

if(in_array('oto', $_POST['select']))
{
    $scrap = new Webscraping();
    /*$doto = $scrap-
>getDetikArticles("https://oto.detik.com/indeks?date=$b1n%2F$t1
%2F$thn"); //detik oto
    for ($i=0; $i <count($doto) ; $i++) {
        $full[] = $doto[$i];
        $arr = explode("\n", $doto[$i]);
        $judul[] = $arr[0];
        $isi[] = $arr[1];
    }*/
    $koto = $scrap-
>getKompasArticles("http://otomotif.kompas.com/search/$thn-$b1n-
$t1"); //kompas otomotif
    for ($i=0; $i <count($koto) ; $i++) {
        $full[] = $koto[$i];
        $arr = explode("\n", $koto[$i]);
        $judul[] = $arr[0];
        $isi[] = $arr[1];
    }
}

if(in_array('tekn', $_POST['select']))
{
    $scrap = new Webscraping();
    /*$dtek = $scrap-
>getDetikArticles("https://inet.detik.com/main/indeks?date=$b1n
%2F$t1%2F$thn"); //detik inet
    for ($i=0; $i <count($dtek) ; $i++) {
        $full[] = $dtek[$i];
        $arr = explode("\n", $dtek[$i]);
        $judul[] = $arr[0];
        $isi[] = $arr[1];
    }*/
    $ktek = $scrap-
>getKompasArticles("http://tekn.kompas.com/search/$thn-$b1n-
$t1"); //kompas tekno
    for ($i=0; $i <count($ktek) ; $i++) {
        $full[] = $ktek[$i];
        $arr = explode("\n", $ktek[$i]);
        $judul[] = $arr[0];
        $isi[] = $arr[1];
    }
}

```

```

    }
}

if(in_array('travel', $_POST['select']))
{
    $scrap = new Webscraping();
    /*$dtra = $scrap-
>getDetikArticles("https://travel.detik.com/indeks?date=$bln
%2F$stgl%2F$thn"); //detik travel
    for ($i=0; $i <count($dtra) ; $i++) {
        $full[] = $dtra[$i];
        $sarr = explode("\n", $dtra[$i]);
        $judul[] = $sarr[0];
        $isi[] = $sarr[1];
    }*/
    $ktra = $scrap-
>getKompasArticles("http://travel.kompas.com/search/$thn-$bln-
$stgl"); //kompas travel
    for ($i=0; $i <count($ktra) ; $i++) {
        $full[] = $ktra[$i];
        $sarr = explode("\n", $ktra[$i]);
        $judul[] = $sarr[0];
        $isi[] = $sarr[1];
    }
}
}

```

### USER NEWS ARTICLE TEXT PREPROCESSING

```

$token = new Tokenization();
$sword = new Stopword();
$stem = new Stemmer();

$cntk = count($k); //count user article

for($i=0;$i<$cntk;$i++)
{
    $tokk = $token->tokenize($k[$i]);
    $swk[] = $sword->removal($tokk);
    //$smk[] = $stem->checkWord($swk);
}

```

### SAVE WORDS FROM USER NEWS ARTICLES TO BAG-OF-WORDS

```

for($j=0;$j<count($swk);$j++)
{
    for($k=0;$k<count($swk[$j]);$k++)
    {
        $db[] = $swk[$j][$k];
        //$db[] = $smk[$j][$k];
    }
}
sort($db);
$dbb = array_values(array_unique($db));

```

### COUNT TERM FREQUENCY(TF)

```
$weight = new TermWeight();
for($l=0;$l<$cntk;$l++)
{
    $ctfk = $weight->countTF($dbs,$swk[$l]);
    //$ctfk = $weight->countTF($dbs,$smk[$l]);
    $tfk[$l] = $ctfk;
}

```

### ONLINE NEWS ARTICLE TEXT PREPROCESSING AND COUNT TERM FREQUENCY

```
$cnt = count($isi);
for($m=0;$m<$cnt;$m++)
{
    $tok = $token->tokenize($isi[$m]);
    $sw = $sword->removal($tok);
    //$sm = $stem->checkWord($sw);
    $ctf = $weight->countTF($dbs,$sw);
    //$ctf = $weight->countTF($dbs,$sm);
    $tf[$m] = $ctf;
}

```

### PUSH USER ARTICLES TF DATA TO ONLINE NEWS ARTICLES

#### ARRAY

```
for($n=0;$n<count($tfk);$n++)
{
    array_push($tf,$tfk[$n]);
}

```

#### COUNT DF, IDF AND TF-IDF

```
$df = $weight->countDF($db,$tf);
$idf = $weight->countIDF(count($tf),$df);
for($o=0;$o<count($tf);$o++)
{
    $keys = array_keys($tf[$o]);
    for($p=0;$p<count($keys);$p++)
    {
        $tf_idf[$o][$keys[$p]] = $tf[$o]
[$keys[$p]]*$idf[$keys[$p]];
    }
}

```

#### MOVE USER ARTICLES TF-IDF RESULT TO NEW ARRAY

```
$cntt = count($tf_idf);
for($a=$cntt;$a>($cntt-$cntk);$a--)
{
    $c[]=$tf_idf[$a-1];
}

```

```

        unset($tf_idf[$a-1]);
    }
    $rev = array_reverse($c);

```

### K-MEANS

```

$cluster = new Cluster();
$skmeans = $cluster->KMeans($tf_idf,$rev);
$relation = $cluster->getRelation();

for($g=0;$g<count($relation);$g++)
{
    echo "<br /><b>Kelompok " . ($g+1) . " - ".str_replace(".txt",
    "", urldecode($filex['name'][$g]))."/><br />";
    if(isset($relation[$g]))
    {
        $keyss = array_keys($relation[$g]);
        for($h=0;$h<count($keyss);$h++)
        {
            $title = $judul[$relation[$g][$h]];
            echo "<form method='post' action='readnews.php'
            target='_blank'>". $title . "<input type='hidden' name='input_name'
            value='".base64_encode(serialize($full[$relation[$g][$h]]))."
            /><input type='submit' value='Baca' /></form>";
        }
    }
    else
    {
        echo "Tidak ada kelompok<br />";
    }
}

```

### SHOW NEWS TITLE AND CONTENT

```

<?
$passed_array = unserialize(base64_decode($_POST['input_name']));
$cek = explode("\n", $passed_array);
?>
<!DOCTYPE html>
<html>
<head>

    <title><?php echo $cek[0];?></title>
</head>
<body>
<?php
    echo $cek[1];
?>
</body>
</html>

```

### CURL

```

function get_data($url) {
    $ch = curl_init();
    $timeout = 5;

```

```

    curl_setopt($ch, CURLOPT_URL, $url);
    curl_setopt($ch, CURLOPT_HEADER, false);
    curl_setopt($ch, CURLOPT_RETURNTRANSFER, true);
    curl_setopt($ch, CURLOPT_MAXREDIRS, 10);
    curl_setopt($ch, CURLOPT_FOLLOWLOCATION, true);
    curl_setopt($ch, CURLOPT_CONNECTTIMEOUT, $timeout);
    curl_setopt($ch, CURLOPT_ENCODING, "");
    curl_setopt($ch, CURLOPT_HTTP_VERSION,
CURL_HTTP_VERSION_1_1);
    $data = curl_exec($ch);
    curl_close($ch);
    return $data;
}

```

### GET NEWS FROM KOMPAS.COM

```

private $content = [];
function getKompasArticles($url)
{
    $html = new simple_html_dom();
    $curl = new cURL();
    $html->load($curl->get_data($url));

    if($html=="")
    {
        echo "Tidak dapat mengambil berita";
    }
    else
    {
        $items = $html->find('div[class=latest--news]',0)-
>find('a');

        $paging = $html->find('div[class=paging]',0);
        if($paging)
        {
            $next = $html->find('a[rel=next]',0);
            if($next)
            {
                for($i=0;$i<count($items);$i++)
                {
                    $html->load($curl-
>get_data($items[$i]->href));
                    $judul = $html-
>find('h1[class=read__title]',0);
                    $judull = strip_tags($judul);

                    $isi = $html-
>find('div[class=read__content]',0);

                    array_push($this->content,
$judull."\n".$isi);
                }
            }
        }
    }
}

```



```

        $this->getKompasArticles($next->href);
    }
    else
    {
        for($i=0;$i<count($items);$i++)
        {
            $html->load($curl-
>get_data($items[$i]->href));

            $judul          =          $html-
>find('h1[class=read__title]',0);
            $judull = strip_tags($judul);

            $isi            =          $html-
>find('div[class=read__content]',0);

            array_push($this->content,
$judull."\n".$isi);
        }
    }
    else
    {
        for($i=0;$i<count($items);$i++)
        {
            $html->load($curl->get_data($items[$i]-
>href));

            $judul          =          $html-
>find('h1[class=read__title]',0);
            $judull = strip_tags($judul);

            $isi            =          $html-
>find('div[class=read__content]',0);

            array_push($this->content, $judull."\n".
$isi);
        }
    }
    return $this->content;
}
}
}

```

### TOKENIZATION CLASS

```

function tokenize($isi)
{
    $html    =    preg_replace("/(<script>(.*)?(\/script>)|(\
(Baca\:&nbsp;.*?\))/s", "", $isi);
    $strip = strip_tags($html);
    $char = array(" ", ".", "-", ":", "/", "(", ")", "\"", "'",
    ",", "!", "&nbsp;", "com");
}

```

```

    $charemove = str_replace($char, " ", $strip);
    $lowercase = strtolower($charemove);
    return $lowercase;
}

```

### STOPWORD REMOVAL CLASS

```

function __construct()
{
    $sword = [];
    $lines = file('class/stopword/id.stopwords.02.01.2016.txt');
    $trim = array_map('trim', $lines);
    array_push($sword, $trim);
    return $sword;
}

function removal($isi)
{
    list($print) = $this->__construct();
    $scnt = str_word_count($isi, 1);
    $strim = array_map('trim', $scnt);
    $srem = array_values(array_diff($strim, $print));
    return $srem;
}

```

### STEMMER CLASS

```

set_time_limit(0);
$word = "";
$prefix = [];
$counter = 0;
$vocal = "aiueo";
$consonant = "bcdfghjklmnpqrstvwxyz";
$derivationprefix = [
    "", ["/^((be)[a-z]/", "/^(ber)/", //1
    "", "/^(be)/", "r"], //2
    er)"/", "/^(ber)/", ""], //3
    ["/^((ber)[consonant][a-z](er)
    [$vocal][a-z]/", "/^(ber)/", ""], //4
    "", //5
    ["/^((be)[bcdfghjklmnpqrstvwxyz]
    (er)[bcdfghjklmnpqrstvwxyz][a-z]/", "/^(be)/", ""], //6
    "", "/^(ter)/", ""], //7
    ["/^((ter)[consonant](er)
    [$vocal][a-z]/", "/^(ter)/", ""], //8
    z)"/", "/^(ter)/", ""], //9
    ["/^((te)[consonant](er)
    [bcdfghjklmnpqrstvwxyz][a-z]/", "/^(te)/", ""], //10
    "", "/^(me)[lrwy][$vocal][a-z]/",

```

```

    "/^(mem)/", "", //11 ["^(mem)[bfv][a-z]/",
    "/^(mem)/", "", //12 ["^(mempe)[rl][a-z]/",
    z]/", "/^(me)/", "", "/^(me)/", "p"], //13 ["^(mem)[r$voCal|$voCal][a-
    "/^(men)/", "", //14 ["^(men)[cdjz][a-z]/",
    "/^(me)/", "", "/^(men)/", "t"], //15 ["^(men)[$voCal][a-z]/",
    "/^(meng)/", "", //16 ["^(meng)[ghq][a-z]/",
    "/^(meng)/", "", "/^(meng)/", "k"], //17 ["^(meng)[$voCal][a-z]/",
    "/^(meny)/", "s"], //18 ["^(meny)[$voCal][a-z]/",
    "/^(mem)/", "p"], //19 ["^(memp)[aiuo][a-z]/",
    "/^(pe)/", "", //20 ["^(pe)[wy][$voCal][a-z]/",
    "/^(per)/", "", "/^(pe)/", "r"], //21 ["^(per)[$voCal][a-z]/",
    "/^(per)/", "", //23 ["^(per)[$consonant][a-z]/",
    ["$voCal][a-z]/", "/^(per)/", "", //24 ["^(per)[$consonant][a-z](er)
    "/^(pem)/", "", //25 ["^(pem)[bfv][a-z]/",
    z]/", "/^(pe)/", "m", "/^(pe)/", "p"], //26 ["^(pem)[r$voCal|$voCal][a-
    "/^(pen)/", "", //27 ["^(pen)[cdjz][a-z]/",
    "/^(pe)/", "", "/^(pen)/", "t"], //28 ["^(pen)[$voCal][a-z]/",
    "/^(peng)/", "", //29 ["^(peng)[ghq][a-z]/",
    "/^(peng)/", "", "/^(peng)/", "k"], //30 ["^(peng)[$voCal][a-z]/",
    "/^(peny)/", "s"], //31 ["^(peny)[$voCal][a-z]/",
    "/^(pe)/", "", //32 ["^(pel)[$voCal][a-z]/",
    ["$voCal][a-z]/", "/^(pe)[bcdfghjkkpqstvxz]/", "", //33 ["^(pe)[bcdfghjkkpqstvxz](er)
    er)[a-z]/", "/^(pe)/", "" //34 ["^(pe)[bcdfghjkkpqstvxz](?!
];

```

```

class Stemmer
{
    function __construct()
    {
        $root = [];
    }
}

```

```

    $lines = file('class/rootword/kata-dasar-indonesia-
mod.txt');
    $lowcase = array_map('strtolower', $lines);
    $trim = array_map('trim', $lowcase);
    array_push($root,$trim);
    return $root;
}

function checkDict($word)
{
    list($dict) = $this->__construct();
    if(in_array($word,$dict))
    {
        return true;
    }
    else
    {
        return false;
    }
}

function setWord($word)
{
    $GLOBALS['word'] = $word;
}

function getWord()
{
    return $GLOBALS['word'];
}

function setCounter($count)
{
    $GLOBALS['counter'] = $count;
}

function getCounter()
{
    return $GLOBALS['counter'];
}

function removeInflectionalSuffixes($word)
{
    //Inflection Suffixes : "-lah", "-kah", "-ku", "-mu",
"-nya"
    //Particle (P) : "-lah", "-kah", "-tah", "-pun"
    //Possessive Pronoun (PP) : "-ku", "-mu", "-nya"

    $wordd = $word;

    if(preg_match("/[a-z]([lkt]ah|pun|nya)$/", $word))
    {
        $wordd = preg_replace("/([lkt]ah|pun|nya)$/", "",
$word);

```

```

        if(preg_match("/[a-z]([km]u|nya)$/", $wordd))
        {
            $wordd = preg_replace("/([km]u|nya)$/", "",
$wordp);
        }
    }
    $this->removeDerivationSuffixes($wordd);
}

function checkDisallowedPrefixSuffix($word)
{
    /*
    Prefix      Disallowed suffixes
    be-   -i
    di-   -an
    ke-   -i, -kan
    me-   -an
    se-   -i, -kan
    te-   -an
    */
    //be- -i
    if(preg_match("/^(be)[a-z](i)$/", $word))
    {
        return true;
    }
    //di- -an
    else if(preg_match("/^(di)[a-z](an)$/", $word))
    {
        return true;
    }
    //ke- -i, -kan
    else if(preg_match("/^(ke)[a-z](i|kan)$/", $word))
    {
        return true;
    }
    //me- -an
    else if(preg_match("/^(me)[a-z](an)$/", $word))
    {
        return true;
    }
    //se- -i, -kan
    else if(preg_match("/^(se)[a-z](i|kan)$/", $word))
    {
        return true;
    }
    //te- -an
    else if(preg_match("/^(te)[a-z](an)$/", $word))
    {
        return true;
    }
    return false;
}

function removeDerivationSuffixes($word)

```

```

{
    //"-i", "-kan", "-an"
    $this->setWord($word);
    if(preg_match("/[a-z](i|an)$/", $word))
    {
        $wordd = preg_replace("/(i|an)$/", "", $word);
        if($this->checkDict($wordd)) //jika
ditemukan di kamus, algoritma berhenti
        {
            $this->setWord($wordd);
        }
        else
        {
            //Step 4a : If a suffix was removed in
Step 3, then disallowed prefix-suffix combinations are checked
using the list in Table 1. If a match is found, then the algorithm
returns.
            if($this-
>checkDisallowedPrefixSuffix($wordd))
            {
                $this->setWord($wordd);
            }
            else
            {
                //Step 4 attempted
                $ina = $this-
>removeDerivationPrefixes($wordd);
                //Step 3a if fail
                if(!$ina)
                {
                    if(preg_match("/[a-z](kan)$/",
$word))
                    {
                        $wordk = preg_replace("/
(kan)$/", "", $word);
                        $this->setCounter(0);
                        if($this-
>checkDict($wordk))
                        {
                            $this-
>setWord($wordk);
                        }
                        else
                        {
                            //Step 4 re-
attempted
                            $kan = $this-
>removeDerivationPrefixes($wordk);
                            //Step 3b if fail
                            if(!$kan)

```

```

        {
            $this->removeDerivationPrefixes($word);
        }
    }
    else
    {
        $this->removeDerivationPrefixes($word);
    }
}

function removeDerivationPrefixes($word)
{
    //Plain prefix : "te-", "me-", "be-", "pe-"
    //Complex prefix : "di-", "ke-", "se-"

    global $prefix, $derivationprefix;

    if($this->checkDict($wordd))
    {
        $this->setWord($wordd);
        return true;
    }
    else
    {
        //Step 4c : If three prefixes have previously
        //been removed, the algorithm returns.
        if($this->getCounter()<3)
        {
            $this->setWord($word);
            if(preg_match("/^(di|[ks]e)[a-z]/",
                $word))
            {
                $wordd = preg_replace("/^(di|[ks]e)/", "", $word);
                $prefix = "/^(di|[ks]e)[a-z]/";

                if($this->checkDict($wordd))
                {
                    $this->setWord($wordd);
                    return true;
                }
            }
        }
    }
}

```

```

else
{
    $this->setCounter($this-
>getCounter()+1);
    $this-
>removeDerivationPrefixes($wordd);
}
}
if(preg_match("/^([tbp]e)[a-z]/", $word))
{
    for($dp=0;$dp<count($derivationprefix);$dp++)
    {
        if(preg_match($derivationprefix[$dp][0], $word))
        {
            if($word=="pelajar")
            {
                $wordd =
                $this-
            }
            else
            {
                $wordd =
                if($this-
                {
                    $this-
                    //cannot
                    return true;
                }
                else
                {
                    if(isset($derivationprefix[$dp][3]))
                    {
                        echo
                        $wordd
                        = preg_replace($derivationprefix[$dp][3], $derivationprefix[$dp]
[4], $word);
                        echo
                        $wordd."$dp \n";
                    }
                    if($this->checkDict($wordd))
                    {
                        $this->setWord($wordd);
                    }
                }
            }
        }
    }
}

```



```

return true;
}
else
{
/array_push($prefix, $derivationprefix[$dp][0]);
$this->setCounter($this->getCounter()+1);
$this->removeDerivationPrefixes($wordd);
return false;
}
}
else
{
>setCounter($this->getCounter()+1);
>removeDerivationPrefixes($wordd);
return
false;
}
}
}
}
}
}
}
return false;
}
function checkWord($word)
{
for($i=0;$i<count($word);$i++)
{
$this->setWord("");
$this->setCounter(0);
unset($prefix);

if(!$this->checkDict($word[$i]))
{
if(strlen($word[$i])>2)
{
$infsu = $this-
>removeInflectionalSuffixes($word[$i]);
$word[$i] = $this->getWord();
}
}
}
return $word;
}
}

```

```
}
```

### TERM WEIGHTING CLASS

```
function countTF($k,$words)
```

```
{
```

```
    $tf = [];
```

```
    sort($k);
```

```
    sort($words);
```

```
    for($i=0;$i<count($k);$i++)
```

```
    {
```

```
        $count_tf = 0;
```

```
        for($j=0;$j<count($words);$j++)
```

```
        {
```

```
            if($words[$j]==$k[$i])
```

```
            {
```

```
                $count_tf++;
```

```
            }
```

```
        }  
        $tf[$k[$i]] = $count_tf;
```

```
        // $tf[$k[$i]] = $count_tf/count($words);
```

```
    }
```

```
    return $tf;
```

```
}
```

```
function countDF($k,$tf)
```

```
{
```

```
    $df = [];
```

```
    for($i=0;$i<count($k);$i++)
```

```
    {
```

```
        $count_df = 0;
```

```
        for($j=0;$j<count($tf);$j++)
```

```
        {
```

```
            $keys = array_keys($tf[$j]);
```

```
            for($l=0;$l<count($keys);$l++)
```

```
            {
```

```
                if($k[$i]==$keys[$l])
```

```
                {
```

```
                    if($tf[$j][$keys[$l]]>0)
```

```
                    {
```

```
                        $count_df++;
```

```
                    }
```

```
                }
```

```
            }
```

```
            $df[$k[$i]] = $count_df;
```

```
        }
```

```
    }
```

```
    return $df;
```

```
}
```

```
function countIDF($total,$df)
```

```
{
```

```
    $idf = [];
```

```

    $keys = array_keys($df);

    for($i=0;$i<count($keys);$i++)
    {
        $idf[$keys[$i]] = log($total/$df[$keys[$i]],10);
    }
    return $idf;
}

```

### K-MEANS CLASS

```
private $cluster, $euclidean, $relate;
```

```

function KMeans($data,$c)
{
    $this->cluster = $c;

    $sed = $this->findEuclideanDistance($data,$c);
    $nearest = $this->findNearestDistance($sed,$data);
    $newc = $this->findNewCentroid($nearest,$data);

    $this->euclidean = $sed;
    $this->relate = $nearest;

    if($newc!=$c)
    {
        $this->KMeans($data,$newc);
    }
}

function getCluster()
{
    return $this->cluster;
}

function getEuclidean()
{
    return $this->euclidean;
}

function getRelation()
{
    return $this->relate;
}

function findEuclideanDistance($data,$c)
{
    for($i=0;$i<count($c);$i++)
    {
        if(isset($c[$i]))
        {
            $keys = array_keys($c[$i]);
            for($k=0;$k<count($data);$k++)
            {

```

```

        $pow = 0;
        for($j=0;$j<count($keys);$j++)
        {
            $pow += ($data[$k][$keys[$j]]-$c[$i]
[$keys[$j]])**2;
        }
        $cluster[$k] = sqrt($pow);
    }
    $sed[$i] = $cluster;
}
else
{
    $c[$i] = null;
}
}
return $sed;
}

function findNearestDistance($sed,$data)
{
    for($a=0;$a<count($data);$a++)
    {
        $column = array_column($sed, $a);
        $min = array_keys($column, min($column));
        $neard[$min[0]][]= $a;
    }
    ksort($neard);
    for($i=0;$i<count($neard);$i++)
    {
        if(!isset($neard[$i]))
        {
            $neard[$i] = null;
        }
        ksort($neard);
    }
    return $neard;
}

function findNewCentroid($neard,$data)
{
    $key = [];
    $key = array_keys($data[0]);

    for($i=0;$i<count($neard);$i++)
    {
        if(isset($neard[$i]))
        {
            $count = count($neard[$i]);
            for($k=0;$k<count($key);$k++)
            {
                $sum = 0;
                for($j=0;$j<$count;$j++)
                {

```

```

                                $sum    +=    $data[$neard[$i][$j]]
[$key[$k]];
                                }
                                $newc[$i][$key[$k]] = $sum/$count;
                                }
                                }
else
{
    for($k=0;$k<count($key);$k++)
    {
        $newc[$i][$key[$k]] = null;
    }
}
}
return $newc;
}

```

**EUCLIDEAN DISTANCE BETWEEN SELECTED NEWS WITH THE FIRST USER NEWS ARTICLE**

$$\begin{aligned}
 &\sqrt{(0-0)^2 + (0-0.97772360528885)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2} \\
 &+ (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0- \\
 &0.80163234623317)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-5.1150144038113)^2 \\
 &+ (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0.97772360528885)^2 \\
 &+ (0-1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + \\
 &(0-0)^2 + (0-0)^2 + (0-0.67669360962487)^2 + (0-1.2787536009528)^2 + \\
 &(0-0.43365556093857)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0- \\
 &1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0- \\
 &2.5575072019057)^2 + (0-1.2787536009528)^2 + (0-0)^2 + (0- \\
 &0.50060235056919)^2 + (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0-0)^2 \\
 &+ (0-0)^2 + (0-1.2787536009528)^2 + (0-1.2787536009528)^2 + (0- \\
 &0.97772360528885)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \\
 &(0-0.67669360962487)^2 + (0.50060235056919-2.0024094022767)^2 + (0- \\
 &0)^2 + (0-0.80163234623317)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0- \\
 &5.6114264236322)^2 + (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0-0)^2 \\
 &+ (0-1.2787536009528)^2 + (0-0.97772360528885)^2 + (0-0)^2 + (0- \\
 &1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \\
 &(0.2053246837943-0.2053246837943)^2 + (0-0)^2 + (0-0.97772360528885)^2 \\
 &+ (0-0.97772360528885)^2 + (0-0.97772360528885)^2 + (0- \\
 &1.6032646924663)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0- \\
 &1.9554472105777)^2 + (0-0)^2 + (0.43365556093857-0.43365556093857)^2 + \\
 &(0-0)^2 + (0-0.97772360528885)^2 + (0-1.7393507898504)^2 + (0-0)^2 + \\
 &(0-0.97772360528885)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \\
 &(0-0)^2 + (0-0)^2 + (0-0)^2 + (0.074633618296904-0.074633618296904)^2 + \\
 &(0-1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0.80163234623317- \\
 &0.80163234623317)^2 + (0-0)^2 + (0-0)^2 + (5.2180523695513-0)^2 + (0-0)^2 \\
 &+ (0.86731112187714-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0- \\
 &2.5575072019057)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0- \\
 &0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \\
 &(0.3245110915135-0.64902218302701)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \\
 &(0-0.67669360962487)^2 + (0-3.8362608028585)^2 + (0-2.5575072019057)^2 \\
 &+ (0-2.5575072019057)^2 + (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + \\
 &(0-0)^2 + (0-0.80163234623317)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 +
 \end{aligned}$$

$$\begin{aligned}
& (0-0)^2 + (0-1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \\
& (0.3245110915135-0)^2 + (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0- \\
& 0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0-1.2787536009528)^2 + (0-0)^2 \\
& + (0-1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0- \\
& 1.2787536009528)^2 + (0-1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0- \\
& 0.80163234623317)^2 + (0-0.57978359661681)^2 + (0.67669360962487-0)^2 \\
& + (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 \\
& + (0-0)^2 + (0-0.97772360528885)^2 + (0-0)^2 + (0-0)^2 + (0- \\
& 1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0.57978359661681-0)^2 \\
& + (0-0.97772360528885)^2 + (0-0)^2 + (0-0)^2 + (0.67669360962487-0)^2 + \\
& (0-1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (2.8989179830841- \\
& 0)^2 + (0-0)^2 + (0-0.67669360962487)^2 + (0-1.2787536009528)^2 + (0- \\
& 0.97772360528885)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0.80163234623317- \\
& 0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0.80163234623317-0)^2 + \\
& (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0-0)^2 + \\
& (0.50060235056919-2.0024094022767)^2 + (0-1.2787536009528)^2 + (0-0)^2 \\
& + (0-0)^2 + (0-0)^2 + (0-0.97772360528885)^2 + (0-0.97772360528885)^2 + \\
& (0-0)^2 + (0-2.5575072019057)^2 + (0-1.2787536009528)^2 = \\
& 14.645963757346
\end{aligned}$$

**EUCLIDEAN DISTANCE BETWEEN SELECTED NEWS WITH THE  
SECOND USER NEWS ARTICLE**

$$\begin{aligned}
& \sqrt{(0-1.2787536009528)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0-0)^2 + \\
& (0-1.2787536009528)^2 + (0-0.67669360962487)^2 + (0-1.2787536009528)^2 \\
& + (0-0)^2 + (0-1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0- \\
& 1.2787536009528)^2 + (0-5.1150144038113)^2 + (0-0.97772360528885)^2 + \\
& (0-0)^2 + (0-0)^2 + (0-1.9554472105777)^2 + (0-1.2787536009528)^2 + (0- \\
& 1.2787536009528)^2 + (0-1.2787536009528)^2 + (0-0.97772360528885)^2 + \\
& (0-0)^2 + (0-0.80163234623317)^2 + (0-1.2787536009528)^2 + (0-0)^2 + \\
& (0-0.97772360528885)^2 + (0-1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + \\
& (0-1.3009666828157)^2 + (0-0)^2 + (0-0)^2 + (0-0.97772360528885)^2 + \\
& (0-0)^2 + (0-1.2787536009528)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0- \\
& 1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 \\
& + (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0- \\
& 1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 \\
& + (0-6.7669360962487)^2 + (0-2.5575072019057)^2 + (0- \\
& 1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0.50060235056919-0)^2 + (0- \\
& 1.2787536009528)^2 + (0-0.80163234623317)^2 + (0-1.2787536009528)^2 + \\
& (0-1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0-6.3937680047641)^2 + (0- \\
& 1.2787536009528)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0-0)^2 + (0-0)^2 \\
& + (0-1.2787536009528)^2 + (0-0)^2 + (0-0.86731112187714)^2 + (0-0)^2 + \\
& (0-0)^2 + (0-0.97772360528885)^2 + (0-1.2787536009528)^2 + \\
& (0.2053246837943-0.10266234189715)^2 + (0-1.2787536009528)^2 + (0-0)^2 \\
& + (0-0.97772360528885)^2 + (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + \\
& (0-0)^2 + (0-0)^2 + (0-2.5575072019057)^2 + (0.43365556093857-0)^2 + \\
& (0-1.2787536009528)^2 + (0-0.97772360528885)^2 + (0-0)^2 + (0- \\
& 1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0- \\
& 1.2787536009528)^2 + (0-0.80163234623317)^2 + (0-1.2787536009528)^2 + \\
& (0-0)^2 + (0-1.2787536009528)^2 + (0.074633618296904- \\
& 0.14926723659381)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0- \\
& 0.97772360528885)^2 + (0.80163234623317-0)^2 + (0-1.2787536009528)^2 + \\
& (0-0)^2 + (5.2180523695513-0)^2 + (0-0)^2 + (0.86731112187714-
\end{aligned}$$



$$\begin{aligned}
& (0.43365556093857-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \\
& (0-0.97772360528885)^2 + (0-2.5575072019057)^2 + (0- \\
& 0.80163234623317)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 \\
& + (0-0)^2 + (0.074633618296904-0.074633618296904)^2 + (0-0)^2 + (0-0)^2 \\
& + (0-0)^2 + (0.80163234623317-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + \\
& (5.2180523695513-3.4787015797009)^2 + (0-1.2787536009528)^2 + \\
& (0.86731112187714-0.43365556093857)^2 + (0-1.2787536009528)^2 + (0- \\
& 0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \\
& (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0-0)^2 + (0- \\
& 1.2787536009528)^2 + (0-0)^2 + (0.3245110915135-0.3245110915135)^2 + \\
& (0-0)^2 + (0-0.97772360528885)^2 + (0-0)^2 + (0-0.67669360962487)^2 + \\
& (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 \\
& + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0- \\
& 0)^2 + (0-1.2787536009528)^2 + (0-1.1595671932336)^2 + \\
& (0.3245110915135-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \\
& (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 \\
& + (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + \\
& (0.67669360962487-2.0300808288746)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + \\
& (0-0)^2 + (0-1.2787536009528)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0- \\
& 1.2787536009528)^2 + (0-0.97772360528885)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 \\
& + (0-0.50060235056919)^2 + (0.57978359661681-1.1595671932336)^2 + \\
& (0-0.97772360528885)^2 + (0-0)^2 + (0-0)^2 + (0.67669360962487-0)^2 + \\
& (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (2.8989179830841- \\
& 1.7393507898504)^2 + (0-0)^2 + (0-0.67669360962487)^2 + (0-0)^2 + (0- \\
& 0.97772360528885)^2 + (0-0)^2 + (0-0.97772360528885)^2 + (0-0)^2 + \\
& (0.80163234623317-0.80163234623317)^2 + (0-2.5575072019057)^2 + (0- \\
& 0)^2 + (0-0)^2 + (0-0)^2 + (0.80163234623317-0)^2 + (0-1.2787536009528)^2 \\
& + (0-1.2787536009528)^2 + (0-0)^2 + (0-1.2787536009528)^2 + \\
& (0.50060235056919-0)^2 + (0-0)^2 + (0-1.2787536009528)^2 + (0-0)^2 + \\
& (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0.97772360528885)^2 + (0-0)^2 + (0-0)^2 = \\
& 11.706900342713
\end{aligned}$$



1647 matches from 115 sources, of which 59 are online sources.

PlagLevel: **8.0%**/59.2%

- [0] (592 matches, 0.0%/59.2%) from a PlagScan document of your organisation...WS\_CLUSTERING.pdf" (
- [1] (60 matches, **5.7%**/8.3%) from a PlagScan document of your organisation...BERITA\_ONLINE.odt" date
- [2] (26 matches, **0.5%**/2.1%) from a PlagScan document of your organisation...AIML\_DATABASE.pdf" dat
- [3] (25 matches, **0.4%**/2.0%) from a PlagScan document of your organisation...hm\_Comparison.odt" date
- [4] (25 matches, **0.4%**/2.0%) from a PlagScan document of your organisation...port\_13020053.pdf" dated
- [5] (24 matches, **0.4%**/1.9%) from your PlagScan document "Devi\_Floren...T'S\_GRADE.pdf" dated 2017-(  
(+ 1 documents with identical matches)
- [7] (24 matches, **0.4%**/1.9%) from your PlagScan document "David\_Kurni...RED\_PROCEDURE.odt" datec  
(+ 1 documents with identical matches)
- [9] (23 matches, **0.4%**/1.9%) from a PlagScan document of your organisation...ject\_13020005.odt" dated
- [10] (23 matches, **0.3%**/1.9%) from your PlagScan document "CHRISTIAN\_A...TUBE\_DATA\_API.odt" date  
(+ 1 documents with identical matches)
- [12] (24 matches, **0.4%**/1.9%) from a PlagScan document of your organisation... \_GEDONG\_SONGO.odt"
- [13] (23 matches, **0.3%**/1.8%) from a PlagScan document of your organisation...R\_ & HISTOGRAM.odt" d
- [14] (22 matches, **0.4%**/1.8%) from a PlagScan document of your organisation...STRATION\_FORM.odt" c
- [15] (22 matches, **0.3%**/1.8%) from your PlagScan document "Vernando\_ A...LTIPLE\_HEAPS..odt" dated  
(+ 1 documents with identical matches)
- [17] (22 matches, **0.3%**/1.8%) from your PlagScan document "Wijayanti\_D...ING\_ARRAYLIST.pdf" dated ;  
(+ 2 documents with identical matches)
- [20] (22 matches, **0.3%**/1.8%) from a PlagScan document of your organisation...REST\_LOCATION.odt" d.
- [21] (22 matches, **0.3%**/1.8%) from a PlagScan document of your organisation...ING\_ARRAYLIST.pdf" da
- [22] (22 matches, **0.3%**/1.8%) from a PlagScan document of your organisation...g\_and\_k-means.pdf" date
- [23] (21 matches, **0.4%**/1.8%) from a PlagScan document of your organisation...g\_Google\_Maps.odt" dat
- [24] (22 matches, **0.3%**/1.8%) from a PlagScan document of your organisation...MATIC\_PATTERN.pdf" d.
- [25] (22 matches, **0.3%**/1.8%) from your PlagScan document "Skripsi\_12....TRA\_ALGORITHM.odt" dated  
(+ 1 documents with identical matches)
- [27] (20 matches, **0.3%**/1.7%) from your PlagScan document "Henry\_Hardi...D\_APPLICATION.odt" dated
- [28] (20 matches, **0.3%**/1.7%) from a PlagScan document of your organisation...D\_APPLICATION.odt" da
- [29] (20 matches, **0.3%**/1.7%) from a PlagScan document of your organisation...D\_APPLICATION.odt" da
- [30] (20 matches, **0.3%**/1.7%) from a PlagScan document of your organisation...ogle\_Maps\_API.odt" date
- [31] (18 matches, **0.3%**/1.7%) from a PlagScan document of your organisation...att\_Algorithm.pdf" dated ;
- [32] (19 matches, **0.3%**/1.6%) from a PlagScan document of your organisation...g\_Google\_Maps.odt" dat
- [33] (18 matches, **0.3%**/1.6%) from a PlagScan document of your organisation...OLLING\_SYSTEM.odt" d.
- [34] (10 matches, **0.1%**/1.0%) from a PlagScan document of your organisation...Project-v2-1.docx" dated
- [35] (9 matches, **0.0%**/1.0%) from a PlagScan document of your organisation...XY\_USING\_JAVA.pdf" dat
- [36] (8 matches, 0.0%/1.0%) from a PlagScan document of your organisation...\_Service\_Chat.pdf" dated 2
- [37] (9 matches, 0.0%/0.9%) from a PlagScan document of your organisation...N\_COEFFICIENT.pdf" date
- [38] (7 matches, 0.0%/0.8%) from repository.unika.ac.id/5786/1/11.02.0027...g William Agitama - COVER,)
- [39] (7 matches, 0.0%/0.8%) from a PlagScan document of your organisation...Z\_APPLICATION.pdf" date
- [40] (7 matches, **0.1%**/0.7%) from a PlagScan document of your organisation...DAG\_Algorithm.pdf" dated
- [41] (6 matches, **0.7%**/0.8%) from indeks.kompas.com/tag/Menteri-ESDM

- ✓ [42] (5 matches, **0.9%**) from [ekonomi.kompas.com/read/2016/09/24/18492....kartu.kredit.jadi.2.25.persen](http://ekonomi.kompas.com/read/2016/09/24/18492....kartu.kredit.jadi.2.25.persen)
- ✓ [43] (6 matches, **0.0%**/0.7%) from [docplayer.net/47167263-Building-network-with-mikrotik-routerboard.htm](http://docplayer.net/47167263-Building-network-with-mikrotik-routerboard.htm)
- ✓ [44] (6 matches, **0.0%**/0.6%) from <https://core.ac.uk/download/pdf/35392981.pdf>
- ✓ [45] (5 matches, **0.0%**/0.6%) from a PlagScan document of your organisation...ject-13020031.doc" dated 2016-09-24
- ✓ [46] (6 matches, **0.2%**/0.6%) from [repository.unika.ac.id/3505/1/08.02.0025...na](http://repository.unika.ac.id/3505/1/08.02.0025...na) Vickey Fitrianingtyas CO
- ✓ [47] (5 matches, **0.6%**) from [bisniskeuangan.kompas.com/read/2017/07/0...naikan.tarif.listrik.hingga.akhir](http://bisniskeuangan.kompas.com/read/2017/07/0...naikan.tarif.listrik.hingga.akhir)
- ✓ [48] (5 matches, **0.0%**/0.5%) from <https://core.ac.uk/download/pdf/35399932.pdf>
- ✓ [49] (3 matches, **0.0%**/0.4%) from <https://core.ac.uk/download/pdf/35399960.pdf>  
(+ 1 documents with identical matches)
- ✓ [51] (3 matches, **0.5%**) from [bisniskeuangan.kompas.com/read/2014/11/1....Ada.Untungnya.bagi.Indones](http://bisniskeuangan.kompas.com/read/2014/11/1....Ada.Untungnya.bagi.Indones)
- ✓ [52] (3 matches, **0.2%**/0.5%) from [https://home.deib.polimi.it/matteucc/Clustering/tutorial\\_html/kmeans.htm](https://home.deib.polimi.it/matteucc/Clustering/tutorial_html/kmeans.htm)
- ✓ [53] (8 matches, **0.4%**) from [php.net/manual/en/ref.curl.php](http://php.net/manual/en/ref.curl.php)  
(+ 1 documents with identical matches)
- ✓ [55] (3 matches, **0.2%**/0.5%) from [home.deib.polimi.it/matteucc/Clustering/tutorial\\_html/kmeans.html](http://home.deib.polimi.it/matteucc/Clustering/tutorial_html/kmeans.html)
- ✓ [56] (7 matches, **0.2%**/0.5%) from a PlagScan document of your organisation... Mersilia AS.docx" dated 2016-09-24  
(+ 1 documents with identical matches)
- ✓ [58] (3 matches, **0.3%**/0.4%) from [www.majalahanalisa.com/2017/07/pastikan-tak-ada-kenaikan-pln-justri](http://www.majalahanalisa.com/2017/07/pastikan-tak-ada-kenaikan-pln-justri)
- ✓ [59] (3 matches, **0.1%**/0.5%) from <https://www.linkedin.com/pulse/k-means-clustering-using-r-jeffrey-strickland>
- ✓ [60] (5 matches, **0.3%**/0.3%) from a PlagScan document of your organisation...0011 Yonathan.pdf" dated 2016-09-24
- ✓ [61] (5 matches, **0.1%**/0.4%) from <https://arxiv.org/pdf/1502.07938>
- ✓ [62] (4 matches, **0.2%**) from a PlagScan document of your organisation...is-Bernadette.doc" dated 2016-11-04
- ✓ [63] (2 matches, **0.0%**/0.2%) from [www.academia.edu/3287106/Scrabble\\_Cheat\\_Sheet](http://www.academia.edu/3287106/Scrabble_Cheat_Sheet)
- ✓ [64] (4 matches, **0.1%**/0.3%) from [arxiv.org/abs/1502.07938?context=cs.LG](http://arxiv.org/abs/1502.07938?context=cs.LG)
- ✓ [65] (4 matches, **0.1%**/0.3%) from <https://www.semanticscholar.org/paper/Do-Not-Use-K-Means-for-Clustering-a46fefdb64a01d1e6390c8>
- ✓ [66] (2 matches, **0.0%**/0.3%) from [java-ml.sourceforge.net/api/0.1.1/net/sf/javaml/clustering/KMeans.html](http://java-ml.sourceforge.net/api/0.1.1/net/sf/javaml/clustering/KMeans.html)  
(+ 3 documents with identical matches)
- ✓ [70] (2 matches, **0.0%**/0.3%) from [www.cbs.dtu.dk/chipcourse/Lectures/ClusteringPCA\\_2010.pdf](http://www.cbs.dtu.dk/chipcourse/Lectures/ClusteringPCA_2010.pdf)
- ✓ [71] (4 matches, **0.1%**/0.3%) from [www.oalib.com/search?kw=Gajen+Chandra+Sarma&searchField=author](http://www.oalib.com/search?kw=Gajen+Chandra+Sarma&searchField=author)
- ✓ [72] (4 matches, **0.1%**/0.3%) from <https://www.researchgate.net/profile/Chandra-Gajen-Chandra-Sarma/publication/309111111-CoverPage=true&origin=publications>
- ✓ [73] (2 matches, **0.0%**/0.3%) from [docplayer.info/35358255-Penerapan-algoritma-k-means-untuk-clusterin](http://docplayer.info/35358255-Penerapan-algoritma-k-means-untuk-clusterin)
- ✓ [74] (3 matches, **0.1%**/0.2%) from a PlagScan document of your organisation...tic\_Algorithm.pdf" dated 2016-09-24
- ✓ [75] (4 matches, **0.1%**/0.2%) from a PlagScan document of your organisation...09 Hempi Indo.doc" dated 2016-09-24
- ✓ [76] (3 matches, **0.1%**) from [labs.jstor.org/blog/](http://labs.jstor.org/blog/)
- ✓ [77] (2 matches, **0.2%**) from <https://sites.google.com/site/dataclusteringalgorithms/k-means-clustering-alg>
- ✓ [78] (4 matches, **0.1%**/0.3%) from <https://core.ac.uk/display/29514835>
- ✓ [79] (4 matches, **0.1%**/0.2%) from a PlagScan document of your organisation...i Ananda (2).docx" dated 2016-09-24
- ✓ [80] (3 matches, **0.1%**/0.2%) from a PlagScan document of your organisation...yanto Martono.pdf" dated 2016-09-24  
(+ 1 documents with identical matches)
- ✓ [82] (2 matches, **0.1%**/0.2%) from [bisniskeuangan.kompas.com/read/2017/06/1...ntau.pemerintah.untuk.t](http://bisniskeuangan.kompas.com/read/2017/06/1...ntau.pemerintah.untuk.t)
- ✓ [83] (2 matches, **0.1%**/0.2%) from [nasional.kompas.com/read/2017/06/22/1801...owi.tak.ada.kenaikan.ha](http://nasional.kompas.com/read/2017/06/22/1801...owi.tak.ada.kenaikan.ha)
- ✓ [84] (2 matches, **0.0%**/0.3%) from [https://en.wikipedia.org/wiki/K-nearest-neighbor\\_estimator](https://en.wikipedia.org/wiki/K-nearest-neighbor_estimator)
- ✓ [85] (3 matches, **0.1%**/0.1%) from a PlagScan document of your organisation...anduning Rat.docx" dated 2016-09-24
- ✓ [86] (3 matches, **0.1%**/0.2%) from a PlagScan document of your organisation...0012 Stefian.docx" dated 2016-09-24
- ✓ [87] (3 matches, **0.1%**/0.1%) from a PlagScan document of your organisation...;antiplagiasi.pdf" dated 2016-09-24
- ✓ [88] (2 matches, **0.2%**) from [note.taable.com/post/23F6A/Cari-Potensi-Lokal-Bekraf/2b-73539T65341-0-5](http://note.taable.com/post/23F6A/Cari-Potensi-Lokal-Bekraf/2b-73539T65341-0-5)
- ✓ [89] (1 matches, **0.0%**/0.2%) from [www.oalib.com/references/14746902](http://www.oalib.com/references/14746902)
- ✓ [90] (2 matches, **0.1%**/0.2%) from <https://www.mikroskil.ac.id/ejurnal/index.php/jsm/article/viewFile/161/95>
- ✓ [91] (2 matches, **0.0%**/0.2%) from [repository.usu.ac.id/bitstream/123456789/49497/1/Reference.pdf](http://repository.usu.ac.id/bitstream/123456789/49497/1/Reference.pdf)
- ✓ [92] (3 matches, **0.1%**/0.2%) from a PlagScan document of your organisation...yati-12020101.pdf" dated 2016-09-24

- ✓ [93] (2 matches, 0.0%/0.2%) from [www.oalib.com/references/9303745](http://www.oalib.com/references/9303745)
- ✓ [94] (3 matches, **0.1%**/0.2%) from [digilib.its.ac.id/public/ITS-paper-30617-5109100153-paper.pdf](http://digilib.its.ac.id/public/ITS-paper-30617-5109100153-paper.pdf)
- ✓ [95] (2 matches, 0.0%/0.2%) from [docplayer.net/35357594-Comparison-jaccar...with-shared-nearest-neigl](http://docplayer.net/35357594-Comparison-jaccar...with-shared-nearest-neigl)
- ✓ [96] (3 matches, **0.2%**) from <https://stackoverflow.com/questions/3414...ay-values-dynamic-input-fields-us>
- ✓ [97] (3 matches, **0.0%**/0.2%) from [bisniskeuangan.kompas.com/read/2017/06/1...e.2.500.desa.yang.belui](http://bisniskeuangan.kompas.com/read/2017/06/1...e.2.500.desa.yang.belui)
- ✓ [98] (3 matches, **0.2%**) from <https://pastebin.com/3Q6EqXek>
- ✓ [99] (3 matches, **0.0%**/0.2%) from [bisniskeuangan.kompas.com/read/2013/06/1...Aakhirnya.Disahkan..Harq](http://bisniskeuangan.kompas.com/read/2013/06/1...Aakhirnya.Disahkan..Harq)
- ✓ [100] (3 matches, **0.0%**/0.2%) from [warta24.com/kades-harus-optimal-gunakan-dana-desa-tanpa-penyek](http://warta24.com/kades-harus-optimal-gunakan-dana-desa-tanpa-penyek)
- ✓ [101] (3 matches, **0.0%**/0.2%) from [nasional.kompas.com/read/2017/06/14/1421...unya.kost-kostan.dan.r](http://nasional.kompas.com/read/2017/06/14/1421...unya.kost-kostan.dan.r)
- ✓ [102] (1 matches, 0.0%/0.2%) from [docplayer.info/35484375-Perbandingan-met...ah-al-qur-an-dalam-bah](http://docplayer.info/35484375-Perbandingan-met...ah-al-qur-an-dalam-bah)  
(+ 1 documents with identical matches)
- ✓ [104] (1 matches, **0.1%**) from a PlagScan document of your organisation...pus cek bab1.docx" dated 2016
- ✓ [105] (1 matches, **0.1%**) from <https://www.edureka.co/blog/k-means-clustering/>
- ✓ [106] (1 matches, **0.1%**) from <https://www.edureka.co/blog/introduction-to-clustering-in-mahout/>
- ✓ [107] (2 matches, 0.0%/0.2%) from [www.worldcat.org/title/proceedings-of-th...-statistics-and-probability/oc](http://www.worldcat.org/title/proceedings-of-th...-statistics-and-probability/oc)
- ✓ [108] (2 matches, 0.0%/0.1%) from a PlagScan document of your organisation...ARIABLE\_DATA.docx" da
- ✓ [109] (1 matches, **0.1%**) from [bisniskeuangan.kompas.com/read/2017/07/0...gulirkan.program.ikkon.2017](http://bisniskeuangan.kompas.com/read/2017/07/0...gulirkan.program.ikkon.2017)
- ✓ [110] (1 matches, **0.1%**) from [indonesia.shafaqna.com/ID/ID/5349155](http://indonesia.shafaqna.com/ID/ID/5349155)
- ✓ [111] (1 matches, **0.1%**) from [megapolitan.kompas.com/read/2014/11/19/1...no.Kenaikan.Harga.BBM.baq](http://megapolitan.kompas.com/read/2014/11/19/1...no.Kenaikan.Harga.BBM.baq)
- ✓ [112] (1 matches, **0.1%**) from <https://twitter.com/kompascom/status/883567836005769216>
- ✓ [113] (1 matches, **0.1%**) from [indonesia.shafaqna.com/ID/ID/5348918](http://indonesia.shafaqna.com/ID/ID/5348918)
- ✓ [114] (1 matches, **0.1%**) from <https://twitter.com/kompascom/status/883639811159990272>

## Settings

Sensitivity: *Medium*

Bibliography: *Consider text*

Citation detection: *Reduce PlagLevel*

Whitelist: --

Analyzed document

-----1/54-----

Cover  
 PROJECT REPORT  
 K-MEANS ALGORITHM IMPLEMENTATION  
 FOR NEWS CLUSTERING  
 DICKY LARSON  
 13.02.0001  
 Faculty of Computer Science  
 Soegijapranata Catholic University  
 2017

i  
 =====2/54=====

APPROVAL AND RATIFICATION PAGE  
 K-MEANS ALGORITHM IMPLEMENTATION FOR NEWS CLUSTERING  
 by  
 DICKY LARSON – 13.02.0001