

RESPONSI PRAKTIKUM NLP

Nama: Syukron Akbar Hidayatullah

NIM: 21110035

LOAD DATA

```
In [46]:  
  
import nltk  
import string  
import pandas as pd  
from sklearn.metrics import accuracy_score  
from sklearn.metrics.pairwise import cosine_similarity  
from sklearn.feature_extraction.text import CountVectorizer
```

```
In [35]:  
  
nltk.download('punkt')  
  
[nltk_data] Downloading package punkt to /root/nltk_data...  
[nltk_data]   Package punkt is already up-to-date!
```

Out[35]:

True

```
In [36]:  
  
file_path = 'data_validation.json'  
with open(file_path, 'r', encoding='utf-8') as file:  
    json_data = file.read()  
  
df = pd.read_json(json_data)  
df.head()
```

Out[36]:

	context	qas
0	Hai! ITTPBot di sini. Ada yang ingin ditanya...	[[{'id': '00001', 'is_impossible': False, 'ques...
1	Kampus Merdeka adalah kebijakan yang dikeluark...	[[{'id': '00006', 'is_impossible': False, 'ques...
2	Jenis kegiatan yang tersedia di program Kampus...	[[{'id': '00010', 'is_impossible': False, 'ques...

3	Studi Independen merupakan salah satu program ...	[[{'id': '00012', 'is_impossible': False, 'ques...
4	Mahasiswa dari semua jurusan pada jenjang D3, ...	[[{'id': '00018', 'is_impossible': False, 'ques...

PRE-PROCESSING TEKS

In [37]:

```
def preprocess_text(text):  
    # Tokenization  
    tokens = nltk.word_tokenize(text)  
    # Lowercasing  
    tokens = [word.lower() for word in tokens if word.isalpha()]  
    # Normalisasi  
    normalized_tokens = [word if word not in string.punctuation else word.lower() for word in tokens]  
    # Merging  
    preprocessed_text = ' '.join(normalized_tokens)  
    return preprocessed_text  
  
df['preprocessed_context'] = df['context'].apply(preprocess_text)  
df.head()
```

Out[37]:

	context	gas	preprocessed_context
0	Hai! ITTPBot di sini. Ada yang ingin ditanya...	[[{'id': '00001', 'is_impossible': False, 'ques...	hai ittpbot di sini ada yang ingin ditanyakan ...
1	Kampus Merdeka adalah kebijakan yang dikeluark...	[[{'id': '00006', 'is_impossible': False, 'ques...	kampus merdeka adalah kebijakan yang dikeluark...
2	Jenis kegiatan yang tersedia di program Kampus...	[[{'id': '00010', 'is_impossible': False, 'ques...	jenis kegiatan yang tersedia di program kampus...
3	Studi Independen merupakan salah satu program ...	[[{'id': '00012', 'is_impossible': False, 'ques...	studi independen merupakan salah satu program ...
4	Mahasiswa dari semua jurusan pada jenjang D3, ...	[[{'id': '00018', 'is_impossible': False, 'ques...	mahasiswa dari semua jurusan pada jenjang dan ...

In [38]:

```
corpus = df['preprocessed_context'].tolist()  
corpus
```

Out[38]:

```
['hai ittpbot di sini ada yang ingin ditanyakan halo saya ittpbot salam kenal ya apa yang bisa saya bantu hari ini helo ini ittpbot  
ada yang ingin kamu ketahui',  
 'kampus merdeka adalah kebijakan yang dikeluarkan oleh kemendikbudristek dengan memberikan hak kepada mahasiswa untuk mengambil m  
ata kuliah di luar program studi selama semester dan berkegiatan di luar perguruan tinggi selama semester perguruan tinggi diberik  
an kebebasan untuk menyediakan kegiatan kampus merdeka yang sesuai dengan kebutuhan dan minat mahasiswanya',  
 'jenis kegiatan yang tersedia di program kampus merdeka yaitu magang bersertifikat studi independen kampus mengajar indonesian in  
ternational student mobility awards diema pertukaran mahasiswa merdeka membangun desa kkn tematik proyek kemanusiaan riset atau pe
```

ternational student mobility awards iisma pertukaran mahasiswa merdeka membangun desa kkn tematik proyek kemanusiaan riset atau penelitian dan wirausaha',

'studi independen merupakan salah satu program kampus merdeka kemendikbud yang bertujuan untuk mewujudkan gagasan mahasiswa dalam mengembangkan diri melalui aktivitas di luar perkuliahan dan menambah kompetensi spesifik serta praktis yang akan digunakan dalam dunia industri',

'mahasiswa dari semua jurusan pada jenjang dan dapat mengikuti program studi independen dengan persetujuan dari perguruan tinggi'

'manfaat yang didapatkan yaitu memberikan kesempatan kepada mahasiswa untuk mengasah dan mendapatkan kemampuan pengetahuan dan sikap di dunia industri dengan cara bekerja dan belajar secara langsung dalam proyek atau permasalahan nyata selain manfaat tersebut keuntungan yang didapatkan yaitu konversi hingga sks dapat melakukan program sertifikasi yang diakui kemendikbudristek selama semester biaya partisipasi serta biaya hidup dan sertifikat sertifikasi atau short course',

'magang bersertifikat adalah bagian dari program kampus merdeka yang bertujuan untuk memberikan kesempatan kepada mahasiswa belajar dan mengembangkan diri melalui aktivitas di luar kelas perkuliahan di program magang bersertifikat mahasiswa akan mendapatkan pengalaman kerja di profesi nyata selama semester dengan pembelajaran langsung di tempat kerja mitra magang mahasiswa akan mendapatkan hard skills maupun soft skills yang akan menyiapkan mahasiswa agar lebih mantab untuk memasuki dunia kerja dan karirnya',

'mahasiswa dari semua jurusan pada jenjang dapat mengikuti program magang bersertifikat dengan persetujuan dari perguruan tinggi'

'manfaat yang didapatkan yaitu mendapatkan kemampuan pengetahuan dan sikap di dunia industri dengan cara bekerja dan belajar secara langsung dalam proyek',

'kampus mengajar adalah salah satu program kampus merdeka kemendikbud berbentuk kegiatan pembelajaran oleh mahasiswa yang memiliki minat dalam memperdalam ilmunya serta membantu pemerataan kualitas pendidikan sekolah tempat praktek mengajar dapat berada di lokasi kota maupun di daerah terpencil',

'syarat mengikuti kampus mengajar yaitu mahasiswa aktif semester memiliki ipk minimum memiliki pengalaman mengajar maupun berorganisasi memperoleh surat rekomendasi dari perguruan tinggi serta belum pernah mengikuti program kampus mengajar',

'manfaat yang didapatkan yaitu mendapatkan bantuan biaya hidup serta ukt bagi yang tidak sedang menerima bantuan pemerintah lainnya mengasah kepemimpinan kreativitas pemecahan masalah inovasi serta terlibat langsung dalam pembelajaran di jenjang sd dan smp',

'iisma adalah singkatan dari indonesian international student mobility awards salah satu program kampus merdeka kemendikbud berupa skema beasiswa dari pemerintah indonesia untuk mendanai siswa dalam program mobilitas ke universitas terkemuka di luar negeri',

'syarat iisma yaitu mahasiswa aktif semester hingga maksimal berumur tahun memiliki ipk minimal memiliki score ielts toefl ibt duolingo english test mendapatkan surat rekomendasi perguruan tinggi serta belum pernah mengikuti program iisma maupun beasiswa lpd',

'manfaat yang didapatkan yaitu konversi sks hingga sks networking bantuan biaya pendaftaran pendidikan serta biaya hidup selama semester test pcr karantina visa transport dan asuransi kesehatan',

'pertukaran mahasiswa merdeka merupakan pertukaran pelajar dengan program studi yang sama ke perguruan tinggi berbeda maupun pertukaran pelajar dengan program studi berbeda ke perguruan tinggi berbeda kegiatan pembelajaran dapat secara tatap muka atau daring program ini bertujuan untuk memperkaya pengalaman dan konteks keilmuan mahasiswa',

'persyaratan program pertukaran mahasiswa merdeka yaitu merupakan mahasiswa aktif semester dan memiliki ipk minimal memiliki pengalaman prestasi tingkat tidak pernah dikenakan sanksi akademik maupun serta mendapatkan surat rekomendasi dari perguruan tinggi dan izin orang tua',

'manfaat yang didapatkan yaitu konversi sks hingga sks networking bantuan biaya hidup dan ukt transport ke kampus tujuan dan biaya rapid',

'program bangkit adalah program kesiapan karier yang didesain oleh google untuk memberikan mahasiswa indonesia paparan langsung dengan praktisi industri serta mempersiapkan mahasiswa dengan keterampilan yang relevan untuk karier sukses di perusahaan teknologi terkemuka',

'persyaratannya yaitu mahasiswa aktif atau minimal semester di kampus terakreditasi bagi atau politeknik di indonesia bagi untuk mahasiswa minimal semester memberikan surat komitmen dan surat rekomendasi universitas',

'bangkit menawarkan pilihan jalur pembelajaran yaitu machine learning mobile development dan cloud computing tidak hanya keterampilan teknologi bangkit juga menyediakan pengembangan bahasa inggris dan soft skill yang penting untuk persiapan karir',

'keterampilan penting untuk karier yang sukses di perusahaan teknologi terkemuka peluang untuk menerapkan pembelajaran anda ke proyek skala nasional dunia nyata kesempatan untuk belajar dari pakar dan praktisi industri saat ini persiapan dan voucher gratis untuk ujian sertifikasi aad atau ace atau tensorflow dan kemungkinan mendapatkan dan menerapkan hingga sks ke program gelar universi

tas anda',

'kuliah kerja nyata tematik atau kknt merupakan suatu bentuk pendidikan dengan cara memberikan pengalaman belajar kepada mahasiswa untuk hidup di tengah masyarakat di luar kampus yang secara langsung masyarakat mengidentifikasi potensi dan menangani masalah sehingga diharapkan mampu mengembangkan potensi desa atau daerah dan meramu solusi untuk masalah yang ada di desa kegiatan kknt diharapkan dapat mengasah softskill kemitraan kerjasama tim lintas disiplin keilmuan atau lintas kompetensi dan leadership mahasiswa dalam mengelola program pembangunan di wilayah perdesaan',

'manfaat dari program kknt yaitu membuat mahasiswa mampu melihat potensi desa mengidentifikasi masalah dan mencari solusi untuk meningkatkan potensi dan menjadi desa mandiri membuat mahasiswa dapat mengaplikasikan ilmu yang dimiliki secara kolaboratif bersama dengan pemerintah desa dan unsur masyarakat untuk membangun desa',

'mahasiswa berasal dari program studi yang mengajukan program kampus merdeka dengan bentuk kegiatan kuliah kerja nyata tematik atau kknt mahasiswa aktif yang terdaftar pada pddikti dibuktikan dengan kartu studi mahasiswa atau ksm termasuk dalam mahasiswa kelas reguler alias bukan mahasiswa alih jenjang dan program lainnya pada program studi strata telah lulus dari sedikitnya sks pada program studi asal dan tidak ada nilai e dibuktikan dengan kartu hasil studi memiliki ipk lebih dari sama dengan dibuktikan dengan kartu hasil studi sehat jasmani dan rohani dibuktikan dengan surat keterangan sehat yang masih berlaku serta tidak sedang hamil bagi wanita lulus dari seleksi yang diadakan oleh institut teknologi telkom purwokerto dan atau mitra tujuan dibuktikan dengan loa atau letter of agreement dilakukan secara berkelompok anggota berjumlah sampai mahasiswa per kelompok dengan minimal disiplin keilmuan yang berbeda ketentuan lain dapat diatur oleh perguruan tinggi pelaksana',

'cara untuk mengakses repository tidak perlu login akun repository hanya diperuntukkan untuk dosen semua babnya open akses kecuali bab dan lampiran pada skripsi bab dan lampiran pada laporan untuk mengakses bab yang terkunci bisa menggunakan layanan online kami infonya dapat dilihat di instagram ittplibrary sistem layanan onlinenya menggunakan google meet nanti request skripsi siapa yang ingin dibuka kemudian akan kami share screen dari sini kemudian dari situ bisa di screen shoot atau screen record',

'perpustakaan ittp berlokasi di gedung rektorat lantai layanan online ittp library bisa diakses pada link atau pada hari dan jam kerja yaitu pada hari senin-jumat pukul untuk kunjungan ke perpustakaan secara offline dimulai dari pukul setiap hari senin kecuali jika ada perubahan secara akan diinformasikan pada instagram ittplibrary untuk dapat masuk dan memakai fasilitas kampus perlu dilakukan reservasi kunjungan terlebih dahulu pada web atau dapat scan barcode di pintu depan perpustakaan untuk mendapatkan link reservasinya',

'untuk persyaratan bebas pustaka dapat dilihat pada ketentuan pada syarat bebas pustaka tanda tangan dalam persyaratan bebas pustaka harus basah skbp akan dikirim setelah semua persyaratan sudah lengkap silahkan bisa mengirim email beserta bukti donasi buku ke perpustakaan baru kemudian skbp akan diproses sumbangan buku guna keperluan bebas pustaka syaratnya buku satu buku sesuai jurusan dan satu buku bebas seperti novel komik dsb dengan minimal harga atau minimal tebal buku halaman donasi point bisa digunakan untuk mengganti sumbangan buku ke perpustakaan jadi untuk menyumbang buku bisa dengan cara memberi buku cetak pada aplikasi ittp digital library dan donasi point pada aplikasi ittp digital library untuk donasi point tidak bisa dibayar langsung ke perpustakaan pembayar donasi point hanya bisa dilakukan pada aplikasi ittp digital library silahkan bisa langsung reply pada email sebelumnya jangan dikirim dengan email baru jika akan dikirim dengan email baru mohon untuk dikirim persyaratan yang lainnya juga untuk programnya silahkan bisa di upload pada drive kemudian lampirkan linknya pada email atau boleh mengirim dalam bentuk atau namun draft bebas pustaka yang lain mohon untuk tidak dikirim dalam bentuk atau',

'untuk pengecekan plagiarisme silahkan dokumennya dikirim ke email perpustakaan formatnya dapat dilihat pada link https ketentuan cek plagiat bisa dilihat pada https pastikan nama dokumen sudah benar sesuai format karena dokumen dengan nama yang salah akan kami kembalikan kepada pengirim untuk diedit terlebih dahulu perpustakaan menggunakan software turnitin dan ithenticate untuk pengecekan plagiarisme dokumen yang dikirim oleh sivitas akademika akan otomatis di cek menggunakan turnitin jika ingin dicek menggunakan ithenticate bisa request pada subject email untuk saat ini perpustakaan tidak menggunakan aplikasi plagiarism detector lagi tidak pengecekan plagiarisme hanya bisa dilakukan oleh perpustakaan atau dosen pembimbing jika untuk syarat sidang akademik akan langsung menolak jika hasilnya didapatkan dari luar kampus petugas akan selalu mengecek email yang masuk secara berkala pada hari dan jam kerja jika load sedang tinggi maka hasil yang didapatkan mungkin bisa cukup lama dan sebaliknya namun kami akan mengirim hasil plagiarisme tidak lebih dari jam jika sudah lebih dari itu silahkan bisa hubungi kami melalui whatsapp layanan pengiriman dokumen pada hari sabtu minggu atau pada libur hari besar akan dicek pada hari kerja selanjutnya',

'untuk menggunakan ruang diskusi bisa reservasi terlebih dahulu pada sebelum reservasi mohon untuk melihat jadwalnya terlebih dahulu pada minimal pemakaian untuk orang maksimal orang maksimal waktu peminjaman jam jika lebih dari itu silahkan untuk menggunakan surat peminjaman ruangan untuk ruang audio visual silahkan langsung bisa dipakai jika memang kosong ruang laktasi merupakan ruangan khusus bagi ibu yang menyusui yang mana akan menampung asi ruangan laktasi dapat langsung dipakai jika memang kosong']

VEKTORISASI

In [39]:

```
def train_bow_model(corpus):  
    vectorizer = CountVectorizer()  
    bow_matrix = vectorizer.fit_transform(corpus)  
    return vectorizer, bow_matrix
```

In [40]:

```
vectorizer, bow_matrix = train_bow_model(df['preprocessed_context'])  
  
print(vectorizer.get_feature_names_out())
```

```
['aad' 'aces' 'ace' 'ada' 'adalah' 'agar' 'agreement' 'akademik'  
'akademika' 'akan' 'aktif' 'aktivitas' 'akun' 'alias' 'alih' 'anda'  
'anggota' 'apa' 'aplikasi' 'asal' 'asi' 'asuransi' 'at' 'atau' 'audio'  
'awards' 'bab' 'babnya' 'bagi' 'bagian' 'bahasa' 'bangkit' 'bantu'  
'bantuan' 'barcode' 'baru' 'basah' 'beasiswa' 'bebas' 'bekerja' 'belajar'  
'belum' 'benar' 'bentuk' 'berada' 'berasal' 'berbeda' 'berbentuk'  
'berjumlah' 'berkala' 'berkegiatan' 'berkelompok' 'berlaku' 'berlokasi'  
'berorganisasi' 'bersama' 'bersertifikat' 'bertujuan' 'berumur' 'berupa'  
'besar' 'beserta' 'biaya' 'bisa' 'boleh' 'bukan' 'bukti' 'buku' 'cara'  
'cek' 'cetak' 'cloud' 'computing' 'course' 'cukup' 'daerah' 'dahulu'  
'dalam' 'dan' 'dapat' 'dari' 'daring' 'dengan' 'depan' 'desa' 'detector'  
'development' 'di' 'diadakan' 'diakses' 'diakui' 'diatur' 'dibayar'  
'diberikan' 'dibuka' 'dibuktikan' 'dicek' 'didapatkan' 'didatapkan'  
'didesain' 'diedit' 'digital' 'digunakan' 'diharapkan' 'diinformasikan'  
'dikeluarkan' 'dikenakan' 'dikirim' 'dilakuakn' 'dilakukan' 'dilihat'  
'dimiliki' 'dimulai' 'dipakai' 'diperuntukkan' 'diproses' 'diri'  
'disiplin' 'diskusi' 'ditanyakan' 'dokumen' 'dokumennya' 'donasi' 'dosen'  
'draft' 'drive' 'dsb' 'dunia' 'duolingo' 'email' 'english' 'fasilitas'  
'format' 'formatnya' 'gagasan' 'gedung' 'gelar' 'google' 'gratis' 'guna'  
'hai' 'hak' 'halaman' 'halo' 'hamil' 'hanya' 'hard' 'harga' 'hari'  
'harus' 'hasil' 'hasilnya' 'helo' 'hidup' 'hingga' 'https' 'hubungi'  
'ibt' 'ibu' 'ielts' 'iisma' 'ilmu' 'ilmunya' 'independen' 'indonesia'  
'indonesian' 'industri' 'infonya' 'inggris' 'ingin' 'ini' 'inovasi'  
'instagram' 'institut' 'international' 'ipk' 'ithenticate' 'ittp'  
'ittpbot' 'ittplibrary' 'itu' 'izin' 'jadi' 'jadwalnya' 'jalur' 'jam'  
'jangan' 'jasmani' 'jenis' 'jenjang' 'jika' 'juga' 'jum' 'jurusan' 'kami'  
'kampus' 'kamu' 'karantina' 'karena' 'karier' 'karir' 'karirnya' 'kartu'  
'ke' 'kebebasan' 'kebijakan' 'kebutuhan' 'kecuali' 'kegiatan' 'keilmuan'  
'kelas' 'kelompok' 'kemampuan' 'kemanusiaan' 'kembalikan' 'kemendikbud'  
'kemendikbudristek' 'kemitraan' 'kemudian' 'kemungkinan' 'kenal' 'kepada'  
'kepemimpinan' 'keperluan' 'kerja' 'kerjasama' 'kesehatan' 'kesempatan'  
'kesiapan' 'ketahui' 'ketentuan' 'keterampilan' 'keterangan' 'keuntungan'  
'khusus' 'kkn' 'kknt' 'kolaboratif' 'komik' 'komitmen' 'kompetensi']
```

'konteks' 'konversi' 'kosong' 'kota' 'kreativitas' 'ksm' 'kualitas'
'kuliah' 'kunjungan' 'lagi' 'lain' 'lainnya' 'laktasi' 'lama' 'lampiran'
'lampirkan' 'langsung' 'lantai' 'laporan' 'layanan' 'leadership'
'learning' 'lebih' 'lengkap' 'letter' 'library' 'libur' 'link' 'linknya'
'lintas' 'loa' 'load' 'login' 'lokasi' 'lpdp' 'luar' 'lulus' 'machine'
'magang' 'mahasiswa' 'mahasiswanya' 'maka' 'maksimal' 'mampu' 'mana'
'mandiri' 'manfaat' 'mantab' 'masalah' 'masih' 'masuk' 'masyarakat'
'mata' 'maupun' 'meet' 'melakukan' 'melalui' 'melihat' 'memakai' 'memang'
'memasuki' 'membangun' 'membantu' 'memberi' 'memberikan' 'membuat'
'memiliki' 'memperdalam' 'memperkaya' 'memperoleh' 'mempersiapkan'
'menambah' 'menampung' 'menangani' 'menawarkan' 'mencari' 'mendanai'
'mendapatkan' 'menerapkan' 'menerima' 'mengajar' 'mengajukan' 'mengakses'
'mengambil' 'mengaplikasikan' 'mengasah' 'mengecek' 'mengelola'
'mengembangkan' 'menggunakan' 'mengidentifikasi' 'mengikuti' 'mengirim'
'meningkatkan' 'menjadi' 'menolak' 'menyediakan' 'menyiapkan'
'menyumbang' 'menyusui' 'meramu' 'merdeka' 'merupakan' 'mewujudkan'
'minat' 'minggu' 'minimal' 'minimum' 'mitra' 'mobile' 'mobilitas'
'mobility' 'mohon' 'muka' 'mungkin' 'nama' 'namun' 'nantinya' 'nasional'
'negeri' 'networking' 'nilai' 'novel' 'nyata' 'of' 'offline' 'oleh'
'online' 'onlinenya' 'open' 'orang' 'otomatis' 'pada' 'pakar' 'paparan'
'partisipasi' 'pastikan' 'pcr' 'pddikti' 'pelajar' 'pelaksana' 'peluang'
'pemakaian' 'pembangunan' 'pembayaran' 'pembelajaran' 'pembimbing'
'pemecahan' 'pemerataan' 'pemerintah' 'peminjaman' 'pendaftaran'
'pendidikan' 'penelitian' 'pengalaman' 'pengecekan' 'pengembangan'
'pengetahuan' 'pengganti' 'pengirim' 'pengiriman' 'penting' 'per'
'perdesaan' 'perguruan' 'perkuliahan' 'perlu' 'permasalahan' 'pernah'
'perpus' 'perpustakaan' 'persetujuan' 'persiapan' 'persyaratan'
'persyaratannya' 'persyaratan' 'pertukaran' 'perubahan' 'perusahaan'
'petugas' 'pilihan' 'pintu' 'plagiarism' 'plagiarisme' 'plagiat' 'point'
'politeknik' 'potensi' 'praktek' 'praktis' 'praktisi' 'prestasi'
'profesi' 'program' 'programnya' 'proyek' 'pukul' 'purwokerto' 'pustaka'
'rapid' 'record' 'reguler' 'rekomendasi' 'rektorat' 'relevan' 'reply'
'repository' 'request' 'reservasi' 'reservasinya' 'riset' 'rohani'
'ruang' 'ruangan' 'saat' 'sabtu' 'salah' 'salam' 'sama' 'sampai' 'sanksi'
'satu' 'saya' 'scan' 'score' 'screen' 'sd' 'sebaliknya' 'sebelum'
'sebelumnya' 'secara' 'sedang' 'sedikitnya' 'sehat' 'sehingga' 'sekolah'
'selain' 'selalu' 'selama' 'selanjutnya' 'seleksi' 'semester' 'semua'
'senin' 'seperti' 'serta' 'sertifikasi' 'sertifikat' 'sesuai' 'setelah'
'setiap' 'share' 'shoot' 'short' 'siapa' 'sidang' 'sikap' 'silahkan'
'singkatan' 'sini' 'sistem' 'siswa' 'situ' 'sivitas' 'skala' 'skbp'
'skema' 'skill' 'skills' 'skripsi' 'sks' 'smp' 'soft' 'softskill'
'software' 'solusi' 'spesifik' 'strata' 'student' 'studi' 'suatu'
'subject' 'sudah' 'sukses' 'sumbangan' 'surat' 'syarat' 'syaratnya'
'tahun' 'tanda' 'tangan' 'tatap' 'tebal' 'teknologi' 'telah' 'telebih'
'telkom' 'tematik' 'tempat' 'tengah' 'tensorflow' 'terakreditasi'
'terdaftar' 'terkemuka' 'terkunci' 'terlebih' 'terlibat' 'termasuk'
'terpencil' 'tersebut' 'tersedia' 'test' 'tidak' 'tim' 'tinggi' 'tingkat'
'toefl' 'transport' 'tua' 'tujuan' 'turnitin' 'ujian' 'ukt' 'universitas'
'unsur' 'untuk' 'upload' 'visa' 'visual' 'voucher' 'waktu' 'wanita' 'web'
'whatsapp' 'wilayah' 'wirausaha' 'ya' 'yaitu' 'yang']

```
vectorized_df = pd.DataFrame(bow_matrix.toarray(), columns=vectorizer.get_feature_names_out())
vectorized_df
```

[illegible]

26	aağ	acces	acğ	ada	adalah	agağ	agreement	akademik	akademikə	akan	...	voucher	waktı	wanita	web	whatsapp	wilayah	wirausahağ	yğ	yaitu	yang
27	0	0	0	0	0	0	0	0	0	3	...	0	0	0	0	0	0	0	0	0	2
28	0	0	0	0	0	0	0	1	1	6	...	0	0	0	0	1	0	0	0	0	4
29	0	0	0	0	0	0	0	0	0	1	...	0	1	0	0	0	0	0	0	0	2

30 rows x 580 columns

MODELLING CHATBOT

In [43]:

```
# FUNGSI MODEL CHATBOT
def model(query, bow_matrix, vectorizer, corpus):
    query_bow = vectorizer.transform([query])
    cosine = cosine_similarity(query_bow, bow_matrix).flatten()
    cosine_index = cosine.argsort()[-1]
    cosine_quest = corpus[cosine_index]
    return cosine_quest

# FUNGSI LOOPING PERTANYAAN
def loop():
    print("Selamat datang di Chatbot Kampus Merdeka! Silakan ajukan pertanyaan atau ketik 'done' untuk keluar.")
    vectorizer, bow_matrix = train_bow_model(corpus)
    while True:
        user_question = input("Anda: ")
        if user_question.lower() == 'done':
            print("Terima kasih telah menggunakan Chatbot Kampus Merdeka. Sampai jumpa!")
            break
        cosine_answer = model(user_question, bow_matrix, vectorizer, corpus)
        print("Chatbot: ", cosine_answer)
        print('\n')

loop()
```

Selamat datang di Chatbot Kampus Merdeka! Silakan ajukan pertanyaan atau ketik 'done' untuk keluar.

Anda: hai

Chatbot: hai ittpbot di sini ada yang ingin ditanyakan halo saya ittpbot salam kenal ya apa yang bisa saya bantu hari ini helo ini ittpbot ada yang ingin kamu ketahui

Anda: Jelaskan mengenai Kampus Merdeka?

Chatbot: jenis kegiatan yang tersedia di program kampus merdeka yaitu magang bersertifikat studi independen kampus mengajar indonesian international student mobility awards iisma pertukaran mahasiswa merdeka membangun desa kkn tematik proyek kemanusiaan riset atau penelitian dan wirausaha

Anda: Apa saja pilihan program Kampus Merdeka yang bisa diambil mahasiswa?

Chatbot: jenis kegiatan yang tersedia di program kampus merdeka yaitu magang bersertifikat studi independen kampus mengajar indonesian international student mobility awards iisma pertukaran mahasiswa merdeka membangun desa kn tematik proyek kemanusiaan riset atau penelitian dan wirausaha

Anda: Apa itu program Bangkit?

Chatbot: program bangkit adalah program kesiapan karier yang didesain oleh google untuk memberikan mahasiswa indonesia paparan langsung dengan praktisi industri serta mempersiapkan mahasiswa dengan keterampilan yang relevan untuk karier sukses di perusahaan teknologi terkemuka

Anda: Definisi Kuliah Kerja Nyata Tematik?

Chatbot: kuliah kerja nyata tematik atau kknt merupakan suatu bentuk pendidikan dengan cara memberikan pengalaman belajar kepada mahasiswa untuk hidup di tengah masyarakat di luar kampus yang secara langsung masyarakat mengidentifikasi potensi dan menangani masalah sehingga diharapkan mampu mengembangkan potensi desa atau daerah dan meramu solusi untuk masalah yang ada di desa kegiatan kknt diharapkan dapat mengasah softskill kemitraan kerjasama tim lintas disiplin keilmuan atau lintas kompetensi dan leadership mahasiswa dalam mengelola program pembangunan di wilayah perdesaan

Anda: Bagaimana cara megakses skripsi atau laporan di repository?

Chatbot: cara untuk mengakses repository tidak perlu login akun repository hanya diperuntukkan untuk dosen semua babnya open access kecuali bab dan lampiran pada skripsi bab dan lampiran pada laporan untuk mengakses bab yang terkunci bisa menggunakan layanan online kami infonya dapat dilihat di instagram ittplibrary sistem layanan onlinenya menggunakan google meet nanti request skripsi si apa yang ingin dibuka kemudian akan kami share screen dari sini kemudian dari situ bisa di screen shoot atau screen record

Anda: Berada dimana perpustakaan ITTP?

Chatbot: perpustakaan ittp berlokasi di gedung rektorat lantai layanan online ittp library bisa diakses pada link atau pada hari dan jam kerja yaitu pada hari senin jum at pukul untuk kunjungan ke perpustakaan secara offline dimulai dari pukul setiap hari senin kecuali jika ada perubahan secara akan diinformasikan pada instagram ittplibrary untuk dapat masuk dan memakai fasilitas kampus perlu dilakuakn reservasi kunjungan terlebih dahulu pada web atau dapat scan barcode di pintu depan perpustakaan untuk mendapatkan link reservasinya

Anda: done

Terima kasih telah menggunakan Chatbot Kampus Merdeka. Sampai jumpa!

EVALUASI MODEL

In [47]:

```
# DATA EVALUASI
test_data = pd.DataFrame({
    'user_question': ["Apa itu program Bangkit?",
                     "Definisi Kuliah Kerja Nyata Tematik?",
                     "Bagaimana cara megakses skripsi atau laporan di repository?",
                     "Apa benefit dari mengikuti program Pertukaran Mahasiswa Merdeka?",
```

```

        "Apa manfaat mengikuti program IISMA?"]],
        'correct_answer': ["program bangkit adalah program kesiapan karier yang didesain oleh google untuk memberikan mahasiswa indone
sia paparan langsung dengan praktisi industri serta mempersiapkan mahasiswa dengan keterampilan yang relevan untuk karier sukses d
i perusahaan teknologi terkemuka",
        "kuliah kerja nyata tematik atau kknt merupakan suatu bentuk pendidikan dengan cara memberikan pengalaman
belajar kepada mahasiswa untuk hidup di tengah masyarakat di luar kampus yang secara langsung masyarakat mengidentifikasi potensi
dan menangani masalah sehingga diharapkan mampu mengembangkan potensi desa atau daerah dan meramu solusi untuk masalah yang ada di
desa kegiatan kknt diharapkan dapat mengasah softskill kemitraan kerjasama tim lintas disiplin keilmuan atau lintas kompetensi dan
leadership mahasiswa dalam mengelola program pembangunan di wilayah perdesaan",
        "cara untuk mengakses repository tidak perlu login akun repository hanya diperuntukkan untuk dosen semua b
abnya open acces kecuali bab dan lampiran pada skripsi bab dan lampiran pada laporan untuk mengakses bab yang terkunci bisa menggu
nakan layanan online kami infonya dapat dilihat di instagram ittpliblibrary sistem layanan onlinenya menggunakan google meet nanti re
quest skripsi siapa yang ingin dibuka kemudian akan kami share screen dari sini kemudian dari situ bisa di screen shoot atau scree
n record",
        "manfaat yang didapatkan yaitu konversi sks hingga 20 sks networking bantuan biaya hidup dan ukt transport
pulang-pergi ke kampus tujuan dan biaya rapid pulang-pergi",
        "manfaat yang didapatkan yaitu konversi sks hingga 20 sks networking bantuan biaya pendaftaran pendidikan
serta biaya hidup selama 1 semester test PCR karantina visa transport dan asuransi kesehatan"]
    })

```

In [48]:

```

# FUNGSI EVALUASI
def eval(chat, test_data, vectorizer, bow_matrix, corpus):
    pred = [chat(query, bow_matrix, vectorizer, corpus) for query in test_data['user_question']]
    true_ans = test_data['correct_answer'].tolist()
    accuracy = accuracy_score(true_ans, pred)
    return accuracy

accuracy = eval(model, test_data, vectorizer, bow_matrix, corpus)
print(f'Akurasi Chatbot: {accuracy * 100:.2f}%')

```

Akurasi Chatbot: 60.00%

TESTING 10 QUEST

In [56]:

```

test_data = pd.DataFrame({
    'user_question': [
        "Apa manfaat Magang Bersertifikat?",
        "Cara mendaftar Kampus Mengajar?",
        "Syarat Indonesian International Student Mobility Awards (IISMA) apa?",
        "Bagaimana cara memakai ruang diskusi?",
        "Cara akses repository perpustakaan Kampus Merdeka gimana?",
        "Program Studi Independen di Kampus Merdeka?",
        "apa saja tahapan dari pengecekan plagiarisme?",
        "Jenis kegiatan di Kampus Merdeka?",
    ]
})

```

```

        "Apa itu Kuliah Kerja Nyata Tematik (KKNT) di Kampus Merdeka?",
        "Proses pendaftaran Program Bangkit bagaimana ya?"
    ]
})

```

In [57]:

```

# FUNGSI TESTING
def test_chatbot(chatbot_function, test_data, vectorizer, bow_matrix, corpus):
    for query in test_data['user_question']:
        print("Anda:", query)
        chatbot_answer = chatbot_function(query, bow_matrix, vectorizer, corpus)
        print("Chatbot:", chatbot_answer)
        print('\n')

# RUN
test_chatbot(model, test_data, vectorizer, bow_matrix, corpus)

```

Anda: Apa manfaat Magang Bersertifikat?

Chatbot: mahasiswa dari semua jurusan pada jenjang dapat mengikuti program magang bersertifikat dengan persetujuan dari perguruan tinggi

Anda: Cara mendaftar Kampus Mengajar?

Chatbot: syarat mengikuti kampus mengajar yaitu mahasiswa aktif semester memiliki ipk minimum memiliki pengalaman mengajar maupun berorganisasi memperoleh surat rekomendasi dari perguruan tinggi serta belum pernah mengikuti program kampus mengajar

Anda: Syarat Indonesian International Student Mobility Awards (IISMA) apa?

Chatbot: iisma adalah singkatan dari indonesian international student mobility awards salah satu program kampus merdeka kemendikbu d berupa skema beasiswa dari pemerintah indonesia untuk mendanai siswa dalam program mobilitas ke universitas terkemuka di luar negeri

Anda: Bagaimana cara memakai ruang diskusi?

Chatbot: untuk menggunakan ruang diskusi bisa reservasi terlebih dahulu pada sebelum reservasi mohon untuk melihat jadwalnya terlebih dahulu pada minimal pemakaian untuk orang maksimal orang maksimal waktu peminjaman jam jika lebih dari itu silahkan untuk menggunakan surat peminjaman ruangan untuk ruang audio visual silahkan langsung bisa dipakai jika memang kosong ruang laktasi merupakan ruangan khusus bagi ibu yang menyusui yang mana akan menampung asi ruangan laktasi dapat langsung dipakai jika memang kosong

Anda: Cara akses repository perpustakaan Kampus Merdeka gimana?

Chatbot: jenis kegiatan yang tersedia di program kampus merdeka yaitu magang bersertifikat studi independen kampus mengajar indonesian international student mobility awards iisma pertukaran mahasiswa merdeka membangun desa kn tematik proyek kemanusiaan riset atau penelitian dan wirausaha

Anda: Program Studi Independen di Kampus Merdeka?

Chatbot: jenis kegiatan yang tersedia di program kampus merdeka yaitu magang bersertifikat studi independen kampus mengajar indonesian international student mobility awards iisma pertukaran mahasiswa merdeka membangun desa kn tematik proyek kemanusiaan riset atau penelitian dan wirausaha

Anda: apa saja tahapan dari pengecekan plagiarisme?

Chatbot: untuk pengecekan plagiarisme silahkan dokumennya dikirim ke email perpustakaan formatnya dapat dilihat pada link [https](https://www.turnitin.com) ke tentuan cek plagiat bisa dilihat pada [https](https://www.turnitin.com) pastikan nama dokumen sudah benar sesuai format karena dokumen dengan nama yang salah akan kami kembalikan kepada pengirim untuk diedit terlebih dahulu perpustakaan menggunakan software turnitin dan ithenticate untuk pengecekan plagiarisme dokumen yang dikirim oleh sivitas akademika akan otomatis di cek menggunakan turnitin jika ingin dicek menggunakan ithenticate bisa request pada subject email untuk saat ini perpustakaan tidak menggunakan aplikasi plagiarism detector lagi tidak pengecekan plagiarisme hanya bisa dilakukan oleh perpustakaan atau dosen pembimbing jika untuk syarat sidang akademik akan langsung menolak jika hasilnya didapatkan dari luar kampus petugas akan selalu mengecek email yang masuk secara berkala pada hari dan jam kerja jika load sedang tinggi maka hasil yang didapatkan mungkin bisa cukup lama dan sebaliknya namun kami akan mengirim hasil plagiarisme tidak lebih dari jam jika sudah lebih dari itu silahkan bisa hubungi kami melalui whatsapp layanan pengiriman dokumen pada hari sabtu minggu atau pada libur hari besar akan dicek pada hari kerja selanjutnya

Anda: Jenis kegiatan di Kampus Merdeka?

Chatbot: jenis kegiatan yang tersedia di program kampus merdeka yaitu magang bersertifikat studi independen kampus mengajar indonesian international student mobility awards iisma pertukaran mahasiswa merdeka membangun desa knk tematik proyek kemanusiaan riset atau penelitian dan wirausaha

Anda: Apa itu Kuliah Kerja Nyata Tematik (KKNT) di Kampus Merdeka?

Chatbot: kuliah kerja nyata tematik atau kknt merupakan suatu bentuk pendidikan dengan cara memberikan pengalaman belajar kepada mahasiswa untuk hidup di tengah masyarakat di luar kampus yang secara langsung masyarakat mengidentifikasi potensi dan menangani masalah sehingga diharapkan mampu mengembangkan potensi desa atau daerah dan meramu solusi untuk masalah yang ada di desa kegiatan kknt diharapkan dapat mengasah softskill kemitraan kerjasama tim lintas disiplin keilmuan atau lintas kompetensi dan leadership mahasiswa dalam mengelola program pembangunan di wilayah perdesaan

Anda: Proses pendaftaran Program Bangkit bagaimana ya?

Chatbot: program bangkit adalah program kesiapan karier yang didesain oleh google untuk memberikan mahasiswa indonesia paparan langsung dengan praktisi industri serta mempersiapkan mahasiswa dengan keterampilan yang relevan untuk karier sukses di perusahaan teknologi terkemuka

ANALISIS HASIL TESTING

Dari 10 pertanyaan test yang diajukan kepada bot, hanya 3 pertanyaan yang dijawab tepat oleh bot dan sisanya dijawab kurang tepat akan tetapi masih sedikit relevan dari pertanyaan yang diajukan. Hasil tersebut menunjukkan bahwa chatbot yang dibangun menggunakan pendekatan Bag of Word dan cosine similarity cukup mampu menggeneralisasi pertanyaan yang diajukan dengan cara yang berbeda.