# KONSEP BERT PADA NATURAL LANGUAGE PROCESSING


Sofa Sofiana, S.Kom., M.Kom.

# KONSEP BERT PADA NATURAL LANGUAGE PROCESSING

BERT (Bidirectional Encoder Representations from Transformers) adalah sebuah model bahasa yang revolusioner dalam bidang pemrosesan bahasa alami (Natural Language Processing, NLP). Dikembangkan oleh Google AI pada tahun 2018, BERT telah mengubah paradigma dalam pemahaman konteks dalam teks.

Konsep utama dari BERT adalah kemampuannya untuk memahami konteks secara mendalam dalam teks dengan memanfaatkan arsitektur Transformer yang canggih. Berbeda dengan model-model sebelumnya, BERT dilatih menggunakan pendekatan "self-supervised learning", di mana model belajar dari data yang tidak berlabel secara langsung.

Salah satu fitur utama dari BERT adalah kemampuannya dalam memproses teks secara "bidirectional", artinya model dapat melihat konteks dari kedua arah dalam sebuah kalimat. Hal ini memungkinkan BERT untuk menghasilkan representasi kata yang lebih kaya dan kontekstual.

Selama pra-pelatihan, BERT dilatih pada data teks yang sangat besar, seperti korpus Wikipedia dan buku-buku. Dengan menggunakan teknik masking, BERT belajar untuk memprediksi kata-kata yang disembunyikan dalam teks, sehingga memungkinkan model untuk memahami konteks lebih baik.

Setelah pra-pelatihan, BERT dapat disesuaikan (fine-tuned) untuk tugas-tugas spesifik dalam NLP, seperti klasifikasi teks, analisis sentimen, atau pertanyaan dan jawaban. Dengan memanfaatkan kemampuan representasi kontekstual yang kaya, BERT telah menjadi landasan bagi berbagai aplikasi NLP yang lebih canggih dan presisi.

Secara keseluruhan, konsep BERT telah membuka pintu bagi kemajuan besar dalam pemrosesan bahasa alami, membawa dampak yang signifikan dalam berbagai bidang seperti pencarian informasi, penerjemahan, dan pengolahan teks berbasis manusia dan mesin.

eureka
media aksara
Anggota IKAPI
No. 225/JTE/2021

0858 5343 1992
eurekamediaaksara@gmail.com
Jl. Banjaran RT.20 RW.10
Bojongsari - Purbalingga 53362

ISBN 978-623-120-443-1
9 786231 204431

# KONSEP BERT PADA NATURAL LANGUAGE PROCESSING

Sofa Sofiana, S.Kom., M.Kom

PENERBIT CV.EUREKA MEDIA AKSARA

**KONSEP BERT PADA NATURAL LANGUAGE PROCESSING**

**Penulis** : Sofa Sofiana, S.Kom., M.Kom

**Desain Sampul** : Eri Setiawan

**Tata Letak** : Herlina Sukma

**ISBN** : 978-623-120-443-1

Diterbitkan oleh : **EUREKA MEDIA AKSARA, MARET 2024**
**ANGGOTA IKAPI JAWA TENGAH**
**NO. 225/JTE/2021**

**Redaksi:**
Jalan Banjaran, Desa Banjaran RT 20 RW 10 Kecamatan Bojongsari Kabupaten Purbalingga Telp. 0858-5343-1992

Surel : eurekamediaaksara@gmail.com

Cetakan Pertama : 2024

## KATA PENGANTAR

Tentang bert pada natural language processing, ini merupakan sebuah terobosan besar yang telah mengubah lanskap dalam memahami dan memproses bahasa alami secara komputasional. BERT, yang merupakan singkatan dari Bidirectional Encoder Representations from Transformers, merupakan model bahasa yang diperkenalkan oleh Google AI pada tahun 2018. Konsep ini mendasarkan dirinya pada transformer architecture yang revolusioner, yang memungkinkan pemahaman konteks dalam kedua arah (bi-directional), berbeda dengan pendekatan sebelumnya yang hanya melihat konteks dalam satu arah.

Penerapan BERT dalam natural language processing telah membawa kemajuan signifikan dalam berbagai bidang, termasuk pemrosesan teks, analisis sentimen, terjemahan mesin, dan lainnya. Kemampuannya untuk memahami konteks dalam kalimat dengan lebih baik telah meningkatkan kualitas hasil pemrosesan bahasa alami secara signifikan.

Dalam kata pengantar ini, kami akan menjelajahi konsep BERT, bagaimana ia beroperasi, dan dampaknya yang luas dalam memajukan kemampuan komputasi untuk memahami dan berinteraksi dengan bahasa manusia.

Tangerang Selatan, 19 Februari 2024

Penulis

## DAFTAR ISI

## DAFTAR GAMBAR

## DAFTAR TABEL

# KONSEP BERT PADA NATURAL LANGUAGE PROCESSING

Sofa Sofiana, S.Kom., M.Kom

# BAB 1 THE ARCHITECTURE OF BERT

## A. Tujuan Pembelajaran

Pada pertemuan ini akan dijelaskan pengetahuan dasar tentang definisi The Architecture Of BERT. Anda harus mampu:

1. Mengetahui Arsitektur BERT
2. Memahami Pengembangan Model BERT Bahasa Khusus Domain
3. Memahami Aplikasi BERT dalam Pemrosesan Bahasa Alami

## B. Uraian Materi

| **Tujuan Pembelajaran 1** |
| --- |
| Pengertian Arsitektur BERT |

### 1. Pengertian Arsitektur BERT

BERT *(Bidirectional Encoder Representations from Transformers)* adalah metode perkembangan dari Transformer yang dikeluarkan pada tahun 2019. Sesuai dengan namanya, BERT hanya melakukan encode dan menghasilkan sebuah model bahasa. Sehingga tidak seperti Transformer, BERT hanya memerlukan encoder (Pratama & Romadhony, 2020).

BERT menggunakan encoder dalam transformator sebagai sub-struktur untuk model pre-training untuk tugas-tugas NLP seperti Sentiment Analysis (SA), Question Answering (QA), Text Summarization (TS) Dalam praktiknya, BERT melakukan dua fase dalam prosesnya

# BAB 2 ENCODER STACK

## A. Tujuan Pembelajaran

Pada pertemuan ini akan dijelaskan pengetahuan dasar tentang definisi Encoder Stack. Anda harus mampu :

1. Mengetahui Encoder Stack

## B. Uraian Materi

| **Tujuan Pembelajaran 1** |
| --- |
| Pengertian Encoder Stack |

Encoder stack adalah struktur data dalam arsitektur Transformer yang terdiridari beberapa blok enkoder yang saling terkait secara urutan. Setiap blokenkoder menerima input dari layer embeding dan pengindalan posisional, dan menggunakan layer Multi-head Attention dan layer Feed-forward untuk mengelola informasi dalam teks. Keluaran dari blok enkoder terakhir digunakan sebagai input untuk setiap blok dekoder dalam stack dekoder. Penggunaan stack encoder dan decoder dalam Transformer memungkinkan model keberlanjutan untuk mengelola informasi dalam teks dengan lebih efisien dan membuat modelyang lebih tahan lama. Meskipun istilah encoder stack juga dapat merujuk padastruktur data lain seperti encoder pada sistem komunikasi atau encoder pada sistem pengkodean video, namun dalam konteks arsitektur Transformer, encoder stack merujuk pada struktur data yang digunakan untuk mengelola informasi dalam teks. Sedangkan ada bagian lain yang

# BAB 3

# FINE-TUNNING BERT DAN HARDWARE CONSTRAINTS

## A. Tujuan Pembelajaran

Pada pertemuan ini akan dijelaskan pengetahuan dasar tentang definisi Fine-Tunning BERT and Hardware Constraints. Anda harus mampu :

1. Mengetahui apa itu Fine-Tuning BERT
2. Mengetahui Hardware Constraints

## B. Uraian Materi

**Tujuan Pembelajaran 1**
Pengertian Fine-Tuning BERT dan Hardware Constraints

### 1. Pengertian Fine-Tuning BERT

BERT adalah model dua arah yang telah dilatih sebelumnya menggunakan sejumlah besar teks dengan tujuan model bahasa bertopeng yang tujuannya adalah untuk memprediksi kata-kata yang disamarkan secara acak dari konteksnya (Grießhaber, Maucher, and Vu 2020). Fine-tuning pada model Bahasa Bert (Bidirectional Encoder Representations from Transformers) merupakan strategi yang memungkinkan penyesuaian model BERT yang sudah dilatih sebelumnya agar lebih sesuai dengan tugas atau data spesifik yang dimiliki pengguna. Ketika menggunakan BERT yang sudah dilatih, kita memanfaatkan pengetahuan luas yang terkandung di dalamnya dari pemrosesan jutaan teks yang berbeda. Fine-tuning memungkinkan model untuk menggabungkan pemahaman yang telah ada dalam BERT

# BAB 4

# INSTALLING THE HUGGING FACE PYTORCH INTERFACE FOR BERT: SPECIFYING CUDA AS THE DEVICE FOR TORCH

## A. Tujuan Pembelajaran

Pertemuan ini akan membahas langkah-langkah instalasi antarmuka PyTorch Hugging Face untuk BERT, dengan penekanan khusus pada penggunaan CUDA sebagai perangkat untuk Torch. Peserta diharapkan dapat:

1. Memahami langkah-langkah instalasi antarmuka PyTorch Hugging Face untuk BERT.
2. Mengerti pentingnya penggunaan CUDA sebagai perangkat untuk meningkatkan kinerja Torch pada pemrosesan bahasa alami.

## B. Uraian Materi

**Tujuan Pembelajaran 1**
Instalasi Antarmuka PyTorch Hugging Face untuk BERT

### 1. Pengantar Instalasi

Langkah pertama dalam memanfaatkan kekuatan BERT adalah menginstal antarmuka PyTorch Hugging Face. Proses ini melibatkan pengunduhan dan konfigurasi lingkungan kerja yang diperlukan untuk memulai penggunaan model BERT. Dengan mengikuti langkah-langkah instalasi ini, Anda akan dapat menyiapkan platform yang sesuai untuk menjalankan model BERT dan mulai menjelajahi kemampuan mumpuni yang ditawarkannya dalam pemrosesan bahasa alami dan tugas-tugas terkait lainnya.

# BAB 5 CREATING SENTENCES, LABEL LISTS, AND ADDING BERT TOKENS

## A. Tujuan Pembelajaran

Pada pertemuan ini akan dijelaskan pengetahuan dasar *(basic science*) tentang definisiCreating sentences, label lists, and adding BERT tokens. Anda harus mampu :

1. Mengetahui Kemampuan BERT dalam membuat kalimat
2. Mengetahui Proses BERT membentuk kalimat

## B. Uraian Materi

| **Tujuan Pembelajaran 1** |
| --- |
| Memahami membuat kalimat, daftar label, dan membuat token BERT |

**1.** Creating Sentences

BERT dapat membuat kalimat dengan cara memproses dan memahami teks yang telah dilatihnya. BERT dilatih pada kumpulan data teks dan kode yang sangatbesar. Kumpulan data ini mencakup berbagai jenis teks, seperti buku, artikel, kode,dan sebagainya. Ketika BERT diminta untuk membuat kalimat, ia akan menggunakan pengetahuannya tentang bahasa yang diperoleh dari kumpulan data pelatihannya. BERT akan mempertimbangkan berbagai faktor saat membuat kalimat, seperti tata bahasa, makna, dan konteks. (Handojo A et al. 2002) << contohreferensi.

# BAB 6 ACTIVATING THE BERTTOKENIZER

## A. Tujuan Pembelajaran

Pada pertemuan ini akan dijelaskan pengetahuan dasar tentang definisi BERTTOKENIZER. Anda harus mampu :

1. Mengetahui Sejarah Tokennizer Bert
2. Memahami membahas aktivasi BertTokenizer dan bagaimana teknik ini dapat meningkatkan kinerja NLP
3. Memahami Aplikasi Tokennizer Bert dalam Pemrosesan Bahasa Alami

## B. Uraian Materi

**Tujuan Pembelajaran 1**
Pengertian Aktivasi Berttokenizer

### 1. Sejarah TokennizerBert

Pemrosesan Bahasa Alami (Natural Language Processing/NLP) adalah bidang yang berkembang pesat dalam komputer dan ilmu komputer. NLP memungkinkan komputer untuk memahami dan memanipulasi bahasa manusia secara otomatis, yang memungkinkan interaksi manusia dengan komputer yang lebih alami dan intuitif. Salah satu teknik yang digunakan dalam NLP adalah tokenisasi, yang merupakan proses membagi teks menjadi bagian-bagian yang lebih kecil yang disebut token. Tokenisasi ini bertujuan untuk memudahkan komputer dalam memahami dan memproses informasi yang terkandung dalam teks.

# BAB 7 PROCESSING THE DATA AND CREATING ATTENTION MASK

## A. Tujuan Pembelajaran

Pada pertemuan ini akan dijelaskan pengetahuan dasar tentang definisi Processing The Data And Creating Attention Mask. Anda harus mampu :

1. Mengetahui Processing The Data And Creating Attention Mask

## B. Uraian Materi

| **Tujuan Pembelajaran 1** |
| --- |
| Pengertian Aktivasi Berttokenizer |

Pentingnya pemrosesan data dan pemanfaatan attention masks dalam meningkatkan pemahamandan analisis data telah menjadi fokus penelitian berbagai studi sebelumnya. Literatur terkait ini memberikan pemahaman mendalam tentang konsep, metode, dan hasil penelitian yang relevan dengan judul "Processing the Data and Creating Attention Masks."

### 1. Pemrosesan Data

Pemrosesan data sebagai langkah awal dalam penelitian ini telah banyak diteliti dalam literatur terkait. Smith et al. (2018) mengeksplorasi teknik pengelolaan data dalam konteks big data analytics, sementara Jones (2020) mengulas strategi pemrosesan data yang efisien untuk meningkatkan kualitas analisis. Penelitian ini merinci berbagai metode dan algoritma yang telah digunakan untuk memproses data sebelum penerapan attention masks.

# BAB 8 PROCESSING THE DATA AND CREATING ATTENTION MASKS

## A. Tujuan Pembelajaran

Pada pertemuan ini akan dijelaskan pengetahuan dasar *(basic science)* tentang defisini Processing The Data and Creating Attention Masks:

1. Penjelasan Tentang Processing The Data and Creating Attention Masks
2. Mengetahui Proses Pengolahan Data
3. Mengetahui Penerapan Studi Kasus Pengolahan data dan pembuatan attention masks
4. Bagaimana Attention Masks Diterapkan Dalam Data yang Sudah Diproses

## B. Uraian Materi

**Tujuan Pembelajaran 1**
Pengertian Processing The Data and Creating Attention Masks

### 1. Pengertian Processing The Data and Creating Attention Masks

Pengolahan data dan pembuatan attention masks adalah dua konsep penting dalam dunia analisis data modern. Pengolahan data melibatkan serangkaingan langkah untuk membersihkan, mengorganisir, dan menganalisis data sehingga dapat diinterpretasikan dengan baik. Attention masks, disisi lain, digunakan untuk memberikan bobot atau penekanan pada bagian tertentu dari data, sambil

# BAB 9 BERT FINE TUNNING

## A. Tujuan Pembelajaran

Pada pertemuan ini akan dijelaskan pengetahuan dasar *(basic science)* tentang defisini Bert Fine Tunning:

1. Penjelasan Tentang Bert Fine Tunning

## B. Uraian Materi

| **Tujuan Pembelajaran 1** |
| --- |
| Pengertian NLP pada Bert Fine Tunning |

Pemrosesan bahasa alami (NLP), seperti analisis sentimen, pertanyaan dan jawaban, dan pengenalan entitas bernama, telah sangat diuntungkan dari kemajuan dalam pembelajaran dalam-dalam dan model berbasis transformer. Di antara model-model ini, BERT (Bidirectional Encoder Representations from Transformers) menonjol karena efektivitasnya dalam menangkap informasi kontekstual dari korpus teks yang besar. Fine-tuning model BERT mengacu pada proses menyesuaikan model pra-pelatihan ini untuk melakukan tugas-tugas NLP tertentu dengan akurasi dan efisiensi tinggi.

### 1. Model BERT Pra-Pelatihan

BERT, diperkenalkan oleh Devlin et al. pada tahun 2018, adalah model berbasis transformer yang dipelatih sebelumnya pada korpus teks besar menggunakan tujuan pemodelan bahasa berantai dan prediksi kalimat berikutnya. Pra-pelatihan BERT menghasilkan representasi kontekstual

# BAB 10 PRETRAINING DAN FINE-TUNING A BERT MODEL

## A. Tujuan Pembelajaran

Pada pertemuan ini akan dijelaskan pengetahuan dasar *(basic science)* tentang defisini Pretraining dan Fine Tuning Bert ?

1. Penjelasan Tentang Pretraining dan Fine Tuning Bert

## B. Uraian Materi

| **Tujuan Pembelajaran 1** |
| --- |
| Pretraining dan Fine Tuning Bert |

Dalam beberapa tahun terakhir, model BERT (Bidirectional Encoder Representations from Transformers) telah menjadi salah satu tonggak penting dalam kemajuan pemrosesan bahasa alami (Natural Language Processing, NLP). Dengan kemampuannya untuk memahami konteks dalam teks dengan mendalam, BERT telah membuka pintu bagi aplikasi NLP yang lebih canggih dan efisien. Namun, untuk memaksimalkan kinerja BERT dalam tugas-tugas spesifik, diperlukan dua tahap penting: pra-pelatihan dan fine-tuning. Makalah ini mengulas proses pra-pelatihan BERT diikuti dengan proses fine-tuning, serta aplikasi dan tantangan yang terkait.

### 1. Pendahuluan

Pra-pelatihan dan fine-tuning adalah dua tahap kunci dalam pengembangan dan penerapan model BERT. Pra-pelatihan BERT melibatkan pembelajaran model pada data teks yang tidak diawasi, sedangkan fine-tuning

4. **Aplikasi dan Manfaat**

   Fine-tuning model BERT telah berhasil diterapkan dalam berbagai tugas NLP, termasuk analisis sentimen, pengenalan entitas bernama, dan banyak lagi. Dengan kemampuannya untuk memahami konteks secara mendalam, BERT memungkinkan aplikasi yang lebih canggih dan presisi dalam memproses teks.

5. **Tantangan dan Pengembangan Masa Depan**

   Meskipun BERT telah menunjukkan hasil yang mengesankan, masih ada sejumlah tantangan yang perlu diatasi, seperti kebutuhan akan dataset yang lebih besar dan peningkatan efisiensi komputasi. Di masa depan, penelitian akan terus fokus pada mengatasi tantangan ini sambil menjelajahi kemungkinan pengembangan model BERT yang lebih lanjut.

## C. Soal

1. Jelaskan apa yang dimaksud dengan pra-pelatihan model BERT.
2. Apa tujuan dari pra-pelatihan dalam konteks model BERT?
3. Sebutkan beberapa tahap utama dalam proses pra-pelatihan model BERT.
4. Apa yang membedakan pra-pelatihan BERT dari fine-tuning BERT?
5. Mengapa perlu menggunakan korpus teks yang besar dalam pra-pelatihan model BERT?
6. Jelaskan konsep "masking" dalam konteks pra-pelatihan BERT.
7. Apa yang dimaksud dengan "bidirectional" dalam nama BERT dan mengapa ini penting dalam konteks pemrosesan bahasa alami?
8. Apa yang dimaksud dengan "self-attention mechanism" dalam arsitektur BERT?
9. Bagaimana model BERT dilatih untuk memprediksi token berikutnya dalam teks?

10. Apa peran dari token [CLS] dalam proses pra-pelatihan BERT?

## D. Kesimpulan

Pra-pelatihan dan fine-tuning merupakan dua tahap penting dalam penggunaan model BERT untuk pemrosesan bahasa alami. Dengan memahami dan mengoptimalkan kedua tahap ini, kita dapat memanfaatkan kekuatan BERT untuk membangun aplikasi NLP yang lebih cerdas dan efisien di masa depan.

## E. Daftar Pustaka

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). RoBERTa: A robustly optimized BERT pretraining approach. arXiv preprint arXiv:1907.11692.

## TENTANG PENULIS

Sofa Sofiana. S.Kom. M. Kom adalah dosen Ilmu Komputer di Universitas Pamulang, Tangerang Selatan. Beliau memperoleh gelar Pascasarjana dari Universitas Bunda Mulia tahun 2013. Sejak tahun 2006, beliau mengajar di beberapa kampus di Jakarta untuk mata kuliah software engineering dan Technopreneurship. Hingga saat ini beliau sebagai konsultan software dan sertifikasi. Sejak tahun 2010 aktif mengajar dan ikut serta di organisasi Ikatan Ahli Informatika Indonesia, Ikatan Asesor Profesional BNSP dan Ikatan Guru Indonesia Tangerang Selatan..