

# FINAL PROJECT OVERVIEW



# NVIDIA-CERTIFIED ASSOCIATE

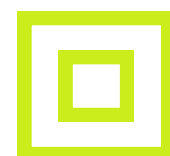
Generative AI LLMS



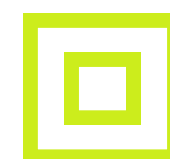
## Objective

Untuk menerapkan pengetahuan dan keterampilan selama *bootcamp* untuk mengembangkan proyek yang memanfaatkan alat dan kerangka kerja NVIDIA untuk Large Language Model (LLM). Proyek tersebut harus menunjukkan keahlian dalam bidang seperti kustomisasi model, rekayasa prompt, analisis data, dan penerapan praktik AI yang dapat dipercaya.

## Group Structure



**Jumlah Tim:** 4 orang per grup



**Maximum Grup per Kelas:** 5

# Project Components

## ☐ 1. Project Proposal

- a. Definisi-kan *problem statement*.
- b. *Outline* Tujuan dan ekspektasi hasil.
- c. Pilih antara menggunakan studi kasus di dalam organisasi atau dataset publik / task

## ☐ 2. Data Collection and Pre-processing

- a. *Source* berdasarkan data yang relevan.
- b. Bersihkan (*clean*) dan *pre-process* data untuk analisa dan modelling.

## ☐ 3. Model Development and Customization

- a. Manfaatkan framework NVIDIA (cth: NVIDIA NeMo) atau kembangkan dan sesuaikan LLMs.
- b. Implemen dan *fine-tune* LLMs sesuai yang diperlukan.

## ☐ 4. Prompt Engineering

- a. Merancang perintah yang efektif untuk mengoptimalkan performa model.

## ☐ 5. Data Analysis and Visualization

- a. Analisis performa model dan wawasan data.
- b. Buat visualisasi untuk menyajikan temuan.

## ☐ 6. Trustworthy AI Considerations

- a. Evaluasi keadilan, bias, dan implikasi etis dari model tersebut.
- b. Menerapkan strategi untuk memastikan model dapat dipercaya

## ☐ 7. Data Analysis and Visualization

- a. Kumpulkan laporan komprehensif yang merinci proses proyek, temuan, dan kesimpulan.
- b. Siapkan presentasi untuk menampilkan proyek di depan kelas

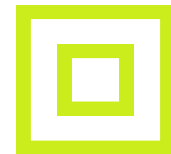


# Guidelines for the Final Project

# Pengajuan Proposal Proyek



**Batas waktu:**  
[XXXX]

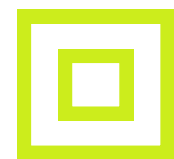


**Isi:**

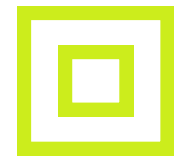
- **Judul:** Ringkasan dan deskriptif.
- **Pernyataan Masalah:** Definisikan dengan jelas masalah yang ingin Anda selesaikan.
- **Tujuan:** Buat daftar tujuan dan apa yang ingin Anda capai.
- **Sumber Data:** Tentukan apakah Anda menggunakan data organisasi atau kumpulan data publik.
- **Metodologi:** Tinjauan singkat tentang pendekatan dan model yang ingin Anda gunakan, menyoroti alat atau kerangka kerja NVIDIA.
- **Hasil yang Diharapkan:** Hasil yang diantisipasi dan potensi dampaknya.



## Eksekusi Proyek



**Alat:** Semua pengembangan harus dilakukan menggunakan Google Colab dan kerangka kerja yang direkomendasikan NVIDIA (misalnya, NVIDIA NeMo).



**Kolaborasi:** Gunakan alat kolaboratif seperti Google Drive, GitHub, atau sistem kontrol versi lainnya untuk mengelola file proyek Anda.



**Dokumentasi:** Simpan dokumentasi yang jelas dan menyeluruh dalam buku catatan dan laporan Colab Anda.

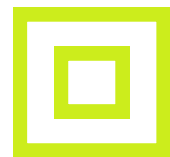
## Hasil Akhir



**Notebook Google Colab:** Notebook komprehensif dengan semua kode, analisis, dan visualisasi menggunakan kerangka kerja NVIDIA jika berlaku.



**Laporan Akhir:** Dokumen terperinci (PDF atau Word) yang mencakup seluruh aspek proyek.



**Presentasi:** *Deck slide* (PowerPoint, Google Slide, dll.) yang merangkum proyek untuk dipresentasikan di depan kelas.





# **Suggested Project Ideas Using Publicly Available Datasets**

(Aligned with NVIDIA Certification)

# Chatbot Khusus untuk Dukungan Pelanggan



## Kumpulan Data:

Customer Support on Twitter Dataset



## Komponen yang Dicakup:

- Pengembangan dan Kustomisasi Model
- Prompt Engineering
- Analisis dan Visualisasi Data
- AI yang dapat dipercaya



## Tugas:

- Memproses terlebih dahulu dan menganalisis tweet dukungan pelanggan.
- Memanfaatkan NVIDIA NeMo untuk membangun dan menyempurnakan model percakapan.
- Menerapkan prompt engineering untuk meningkatkan akurasi dan relevansi respons.
- Evaluasi kinerja chatbot menggunakan metrik seperti akurasi respons dan kepuasan pengguna.
- Pastikan chatbot mematuhi pedoman etika dan meminimalkan bias.

**Alat NVIDIA:** [NVIDIA NeMo](#)

# Rangkuman Teks untuk Makalah Penelitian



## Kumpulan Data:

arXiv Papers Dataset



## Komponen yang Dicakup:

- Pengembangan dan Kustomisasi Model
- Prompt Engineering
- Analisis dan Visualisasi Data
- AI yang dapat dipercaya

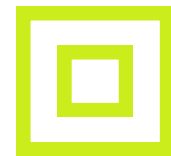


## Tugas:

- Memproses terlebih dahulu makalah penelitian dan abstraknya.
- Menyempurnakan model Transformer terlatih (misalnya, BERT, GPT) menggunakan NVIDIA NeMo untuk tugas peringkasan.
- Petunjuk desain untuk meningkatkan kualitas ringkasan.
- Evaluasi ringkasan menggunakan metrik seperti ROUGE dan BLEU.
- Visualisasikan performa model dan bandingkan dengan model dasar.
- Diskusikan pertimbangan etis dalam peringkasan otomatis.

**Alat NVIDIA:** [NVIDIA NeMo](#)

# Analisis Sentimen untuk Review Produk



## Kumpulan Data:

Amazon Product Reviews Dataset



## Komponen yang Dicakup:

- Pengembangan dan Kustomisasi Model
- Prompt Engineering
- Analisis dan Visualisasi Data
- AI yang dapat dipercaya

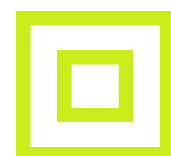


## Tugas:

- Memproses ulang data ulasan produk Amazon.
- Menerapkan dan menyempurnakan LLM menggunakan NVIDIA NeMo untuk klasifikasi sentimen.
- Memanfaatkan rekayasa cepat untuk mengoptimalkan akurasi analisis sentimen.
- Evaluasi performa model menggunakan metrik seperti akurasi, presisi, perolehan, dan skor F1.
- Visualisasikan distribusi sentimen dan wawasan utama.
- Menilai dan memitigasi bias apa pun dalam model analisis sentimen.

**Alat NVIDIA:** [NVIDIA NeMo](#)

# Pembuatan Laporan Otomatis untuk Data Keuangan

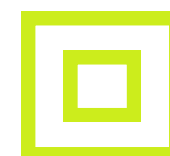


## **Kumpulan Data:** Financial News Dataset



## **Komponen yang Dicakup:**

- Pengembangan dan Kustomisasi Model
- Rekayasa Cepat
- Analisis dan Visualisasi Data
- AI yang dapat dipercaya

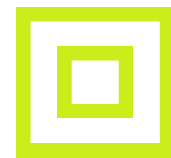


## **Tugas:**

- Memproses awal dan menganalisis kumpulan data berita keuangan.
- Gunakan NVIDIA NeMo untuk menyempurnakan model Transformer untuk pembuatan laporan.
- Desain petunjuk untuk mengekstrak wawasan keuangan spesifik dan menghasilkan narasi.
- Mengevaluasi keakuratan dan relevansi laporan yang dihasilkan.
- Visualisasikan metrik dan tren keuangan utama yang diambil dari laporan.
- Memastikan sistem menghasilkan narasi keuangan yang tidak memihak dan akurat.

**Alat NVIDIA:** [NVIDIA NeMo](#)

# Penciptaan Basis Pengetahuan untuk Informasi Kesehatan



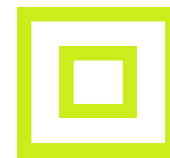
## Kumpulan Data:

Medical Transcription Dataset



## Komponen yang Dicakup:

- Pengembangan dan Kustomisasi Model
- Rekayasa Cepat
- Analisis dan Visualisasi Data
- AI yang dapat dipercaya



## Tugas:

- Proses awal data transkripsi medis.
- Menerapkan dan menyempurnakan LLM menggunakan NVIDIA NeMo untuk ringkasan informasi.
- Memanfaatkan rekayasa cepat untuk memastikan dihasilkannya ringkasan medis yang akurat dan andal.
- Evaluasi performa model menggunakan metrik khusus domain.
- Visualisasikan efektivitas sistem basis pengetahuan.
- Menilai pertimbangan etis dan memastikan kepatuhan terhadap standar data layanan kesehatan.

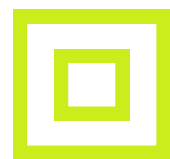
**Alat NVIDIA:** [NVIDIA NeMo](#)





# Langkah-Langkah Eksekusi Proyek

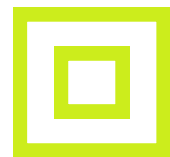
# Membentuk Kelompok dan Memilih Proyek



**Langkah 1:** Bentuklah kelompok yang terdiri dari 4 anggota.

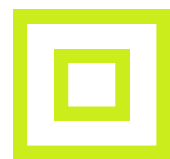


**Langkah 2:** Pilih ide proyek dari daftar yang disarankan atau usulkan ide baru berdasarkan kebutuhan organisasi Anda.

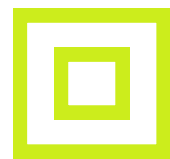


**Langkah 3:** Kirimkan proposal proyek Anda untuk disetujui.

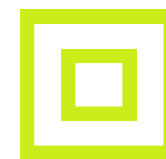
## Pengembangan Proyek



**Penanganan Data:** Kumpulkan dan proses awal kumpulan data yang dipilih.



**Implementasi Model:** Mengembangkan dan menyempurnakan model yang diperlukan menggunakan kerangka kerja NVIDIA (misalnya, NVIDIA NeMo) di Google Colab.



**Prompt Engineering:** Rancang dan sempurnakan perintah untuk mengoptimalkan keluaran model.

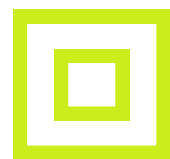


**Analisis:** Melakukan analisis data dan memvisualisasikan hasilnya.



**AI yang Dapat Dipercaya:** Menilai dan mengatasi pertimbangan keadilan, bias, dan etika.

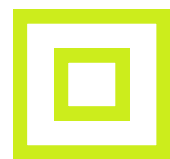
## Dokumentasi dan Pelaporan



**Colab Notebook:** Pastikan notebook Anda terdokumentasi dengan baik dengan penjelasan dan visualisasi kode yang jelas menggunakan alat NVIDIA.

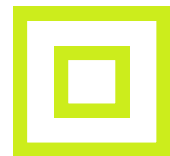


**Laporan Akhir:** Menyusun laporan komprehensif yang mencakup seluruh aspek proyek.



**Presentasi:** Siapkan presentasi singkat untuk menunjukkan tujuan, metodologi, hasil, dan kesimpulan proyek Anda.

# Penyerahan



**Batas waktu:** [xxx]



**Format:** Kirimkan semua kiriman melalui platform yang ditentukan (misalnya LMS, Google Classroom).



# Evaluation Criteria



# Kriteria Evaluasi



## Teknis Implementasi (40%)

- Pemrosesan Awal Data: Efektivitas dan efisiensi.
- Pengembangan dan Kustomisasi Model: Kesesuaian model yang digunakan dan implementasinya menggunakan alat NVIDIA.
- Prompt Engineering: Kreativitas dan efektivitas dalam merancang prompt.
- Evaluasi Kinerja: Akurasi dan relevansi metrik evaluasi.



## Analisis dan Visualisasi Data (20%)

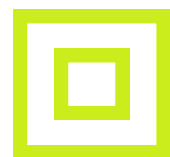
- Insightfulness: Kualitas wawasan data yang diperoleh.
- Kualitas Visualisasi: Kejelasan dan efektivitas visualisasi.

## Kriteria Evaluasi



### **Trustworthy AI Considerations (15%)**

- Penilaian Kewajaran dan Bias: Ketelitian dalam mengevaluasi bias model.
- Implikasi Etis: Memahami dan mengatasi masalah etika.



### **Dokumentasi dan Pelaporan (15%)**

- Kejelasan dan Organisasi: Laporan terstruktur dengan baik dan ditulis dengan jelas.
- Kelengkapan: Cakupan seluruh komponen proyek.



### **Presentasi (10%)**

- Keterlibatan: Kemampuan untuk melibatkan dan memberikan informasi kepada audiens.
- Kejelasan: Penjelasan yang jelas mengenai tujuan, metode, dan temuan proyek.



# Final Project Submission Checklist

# Checklist

Sebelum menyerahkan tugas akhir, pastikan komponen-komponen berikut disertakan:

## 1. Buku Catatan Google Colab:

- Semua sel kode dieksekusi dan hasilnya terlihat.
- Komentar dan penjelasan yang jelas untuk setiap langkah.
- Pengorganisasian dan pemformatan yang tepat menggunakan kerangka kerja NVIDIA.

## 1. Laporan Akhir:

- Pendahuluan: Pernyataan masalah dan tujuan.
- Metodologi: Pengumpulan data, prapemrosesan, pemilihan model, dan rekayasa cepat menggunakan alat NVIDIA.
- Hasil: Performa model, wawasan data, visualisasi.
- AI yang Dapat Dipercaya: Keadilan, penilaian bias,

Slide Presentasi:

- Slide yang jelas dan ringkas yang mencakup semua aspek utama proyek.
- Alat bantu visual untuk mendukung penjelasan.
- Presentasi yang diatur waktu agar sesuai dengan waktu yang ditentukan.

## 1. Dinamika Kelompok:

- Pastikan semua anggota kelompok berkontribusi dan memahami proyek.
- Mencerminkan kolaborasi dan pembagian tugas dalam laporan.
-



# Final Tips for Success

# Checklist

- **Mulai Lebih Awal:** Mulailah mengerjakan proyek sesegera mungkin untuk memberikan waktu yang cukup untuk penelitian, pengembangan, dan pemecahan masalah.
- **Komunikasi yang Jelas:** Pertahankan komunikasi yang teratur dalam kelompok untuk memastikan semua orang memiliki pemahaman yang sama.
- **Pengembangan Berulang:** Kembangkan proyek Anda secara bertahap, terus-menerus menguji dan menyempurnakan setiap komponen.
- **Mencari Umpan Balik:** Jangan ragu untuk mencari umpan balik dari instruktur atau rekan untuk meningkatkan proyek Anda.
- **Fokus pada Kejelasan:** Pastikan laporan akhir dan presentasi Anda menyampaikan dengan jelas tujuan, metode, dan temuan proyek Anda.



# Certification Official Page

<https://www.nvidia.com/en-us/learn/certification/generative-ai-llm-associate/>



**Training**

For Organizations

For Individuals ▾

For Educators

Certification

NVIDIA-Certified Associate

## Generative AI LLMs

(NCA-GENL)

[Register for Exam](#)



**NVIDIA.**

Certified  
**Associate**

**Gen AI LLMs**

