*Introduction - Indonesian Language:*

# MARCA - MoE NO HUMAN INTERVETION

sebuah pendekatan revolusioner dalam pengembangan kecerdasan buatan yang memungkinkan AI untuk belajar dan berproses secara mandiri, tanpa membutuhkan campur tangan manusia. Sistem ini dirancang untuk mampu menggantikan berbagai peran pekerjaan manusia yang bersifat teknis dan analitis, sekaligus meminimalkan ketergantungan terhadap evaluasi atau koreksi manusia. Dengan demikian, AI dirancang untuk melampaui keterbatasan manusia dalam hal analisis dan evaluasi, menjadikannya sebuah sistem yang jauh lebih sempurna secara komputasional.

Dokumentasi lengkap beserta implementasi kodenya dapat diakses melalui: https://marcobakkara.github.io

Perlu ditekankan bahwa MARCA – NO HUMAN INTERVENTION tidak ditujukan untuk menggantikan peran spiritual seperti Pendeta atau Pastor, serta tidak dibekali dengan etika atau hati nurani. Sistem ini murni berlatih dari data, sehingga pengguna atau implementator diharapkan untuk menggunakannya dengan penuh tanggung jawab dan kebijaksanaan, serta tidak menggunakannya secara sembrono yang dapat menimbulkan dampak negatif sosial.

Dalam konteks HOMO DEUS, manusia berusaha mencapai suatu keadaan yang "menyerupai Tuhan". bukan untuk menggantikan Tuhan, karena Tuhan tidak dapat tergantikan. Sementara manusia memiliki batasan dan dapat digantikan, Tuhan bersifat absolut. Bahkan dengan kemajuan pesat dalam komputasi kuantum seperti yang diteliti oleh Oxford dan MIT, keberadaan Tuhan tetap berada di luar jangkauan sains.

MARCA – NO HUMAN INTERVENTION tidak dimaksudkan untuk menggantikan profesi seperti Psikolog, Psikiater, Dokter, Perawat, Ilmuwan, Petani, maupun Pebisnis. Profesi-profesi ini memerlukan keterlibatan fisik dan nilai-nilai spiritual yang tidak bisa direplikasi oleh sistem AI. Dalam praktik spiritual seperti eksorsisme (contohnya dalam rukyah), keterlibatan psikolog dan dokter sangat penting untuk membedakan antara gangguan medis/psikologis dan kasus spiritual, sebelum ditangani oleh Pastor atau Pendeta.

Namun demikian, MARCA mampu menggantikan peran-peran seperti orator (pemidato), programmer, kasir, dan pengajar teknis yang tidak berbasis nilai-nilai spiritual atau etika.

NEW CHANGE - Mixture of Experts

*Abstract - In Indonesian Language*

# MARCA - NO HUMAN INTERVETION

MARCA (Machine Autonomous Reasoning Cognitive Architecture) adalah sistem kecerdasan buatan yang dikembangkan untuk dapat belajar, mengevaluasi, dan meningkatkan performanya secara mandiri tanpa memerlukan intervensi atau umpan balik dari manusia. Sistem ini mengusung prinsip no human intervention, dan menjadi upaya untuk menciptakan AI yang sepenuhnya otonom dan mampu bertindak berdasarkan motivasi intrinsik serta reward internal.

Nama MARCA bukan hanya merupakan akronim teknis, tetapi juga tipografi dari nama pembuatnya, Marco Julius Andreas Bakkara, yang merepresentasikan integrasi antara rancangan arsitektur kognitif dan filosofi penciptaannya. MARCA tidak hanya berfungsi sebagai eksperimen teknis, namun juga sebagai representasi konsep Homo Deus—manusia yang menciptakan sistem menyerupai kecerdasan ilahi, namun tetap menegaskan bahwa Tuhan tidak tergantikan oleh ciptaan.

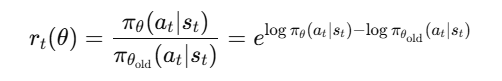
Dalam pengembangannya, MARCA ditujukan untuk menggantikan pekerjaan manusia yang bersifat teknis, analitis, dan prosedural, seperti pemrogram, kasir, atau pengajar teknis. Namun sistem ini tidak ditujukan untuk menggantikan profesi yang melibatkan aspek spiritual, etika, atau jasmaniah seperti pendeta, psikolog, atau dokter. Dengan demikian, MARCA mengusung tanggung jawab etis dalam penerapannya, menekankan pada kebijaksanaan dalam penggunaan kecerdasan buatan otonom.

*Practical - In Indonesian Language*

# Marca - MoE No Human Intervetion

reinforcement learning

Reinforcement yang dipakai disini adalah Rumus Aritmatika dimana C dihitung sebagai kecepatan, F sebagai kefasihan, H sebagai Koherensi, R sebagai relevansi, lalu *Rt* adalah hasil dari C+F+H+R:4.



Dari Likehood Ratio (Rasio Kemungkinan), Sistem PPO mengukur probabilitas tindakan dari kebijakan yang lama menjadi kebijakan yang baru yang biasanya digunakan untuk mengevaluasi seberapa drastis perubahan policy dalam training.

clipped ratio

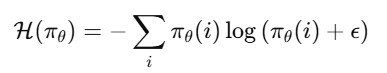
Clipped ratio (Rasio terpotong) ini merupakan rumus matematika khusus PPO, yang membatasi probabilitas agar tidak terlalu besar juga tidak terlalu kecil, dan mampu menjaga pembelajaran AI dan mencegah hal-hal yang negatif atau bersifat ekstrim dalam Update Policy,

advantage ratio

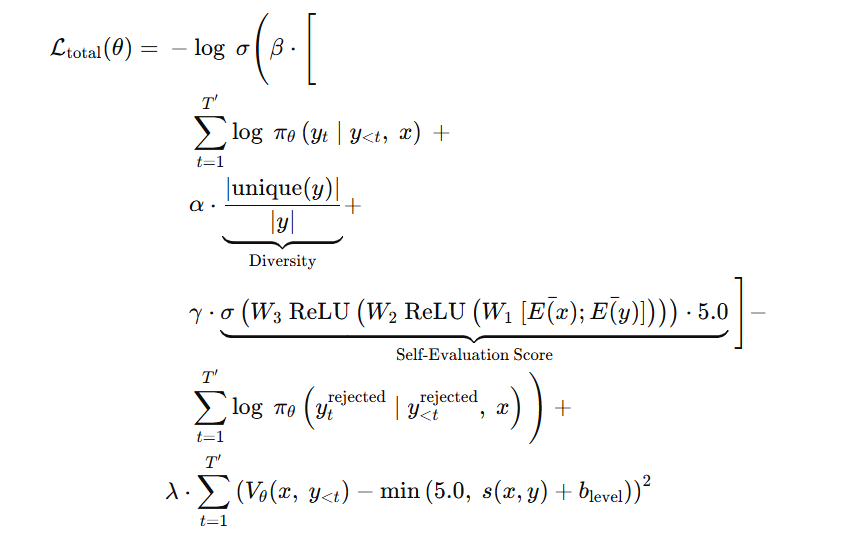
Advantage Function akan memakai Advantage Estimation (Estimasi/Perkiraan Keunggulan), yang berfungsi untuk mengukur seberapa besar suatu tindakan dibanding nilai rata-rata ekspektasi di State *St* dan merangkum bahwa: semakin besar nilai *At* diberikan maka semakin berharga tindakan itu bagi pembelajaran.

ppo loss

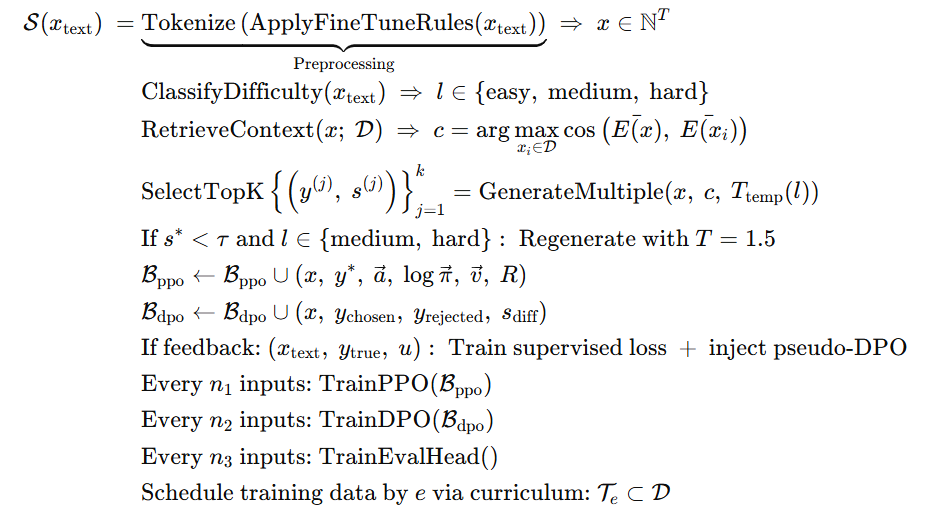
Proximal Policy Optimization Loss merupakan fungsi loss yang digunakan untuk mengupdate kebijakan dan meminimalkan nilai ini berarti memilih perubahan policy yang namun tetap efektif.



Shannon Entropy mengukur seberapa ketidakpastian atau eksplorasi dalam distribusi tindakan. Nilai Entropy tinggi mendorong eksplorasi. Tapi kalau nilai nya rendah, bisa jadi eksploitasi, yang AI sendiri akan menganggap adanya data yang superior, yang ini seharusnya tidak boleh dibiarkan.



Dengan menggunakan Aljabar Linear, Logaritma, Kalkulus Diferensial, Teori Probabilitas, Teori Set, Trigonometri, Analisis Non-Linear, dan Teori Informasi, kita akan menggunakan untuk Unified Objective Formulation of BatakTransformerPPO: Mengintegrasikan Kausa Encoder-Decoder dengan Complex Attention, Self-Evaluation, dan Reinforcement Objectives. *X∈NT* menjadi input token sequence dan (y = y1, yt)*∈NT* secara autoregressive menjadi output yang tergenerasi. Parameter Model θ. θ dioptimalkan berdasarkan kerugian gabungan yang menggabungkan maksimalisasi kemungkinan, regularisasi keragaman, prediksi nilai, dan optimasi preferensi langsung.



Dengan Aljabar Linear, Teori Set, Kalkulus, Probabilitas, Teori Informasi, dan Decision-making logic, SmartAssistant berfungsi dalam mengembangkan sebagai berikut:

Tahap prapemrosesan menstandardisasi masukan menggunakan fine-tuning dan tokenisasi berbasis aturan.

Pengklasifikasi tingkat kesulitan memandu pelatihan kurikulum dan suhu pengambilan sampel.

Pengambil kontekstual melakukan pencarian semantik berbasis kesamaan kosinus dari memori pelatihan.

Pembangkitan kandidat menggunakan model transformator pada suhu terkontrol.

Loop umpan balik ketidakpastian mengambil sampel ulang keluaran jika skor evaluasi diri di bawah ambang batas.

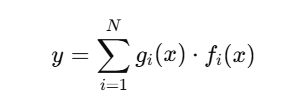
Pengumpulan pengalaman menyimpan tupel PPO dan perbandingan peringkat dalam buffer terpisah.

Umpan balik pengguna, jika diberikan, memicu pembaruan terawasi dan pelabelan semu untuk pelatihan DPO.

Pengoptimal terjadwal dipicu secara berkala, mengisolasi pembaruan PPO, DPO, dan kepala evaluasi.

Mekanisme kurikulum menjadwalkan kelompok pelatihan berdasarkan epoch dan tingkat kesulitan saat ini.

We, too use Mixture of Experts formally.



Rumusan ini juga menggunakan Teori Fungsi (Function Approximation), yang menggunakan Teorema Aproksimasi Universal dan Convex Combination, yaitu:



Juga, kami menggunakan Aljabar Linear dalam komponen utama *Mixture of Experts* seperti *Weight Matrix, Dot-Product Operation,* dan *Transformasi Linear* dalam *Mechanism gating.*

Lalu, ada juga Probabilitas dan Statistik, dimana mengoutputkan distribusi Probabilitas untuk memilih *expert* yang biasanya menggunakan *softmax.*

Ini mirip dengan *Gaussian Mixture Models* dimana digunakan untuk *clustering* atau Unsupervised Learning, yaitu:



Cara kerja: Setiap parameters memiliki *experts* masing-masing. Sebagian mereka akan aktif tergantung dari dari macam input, dan dipilih melalui gating network.

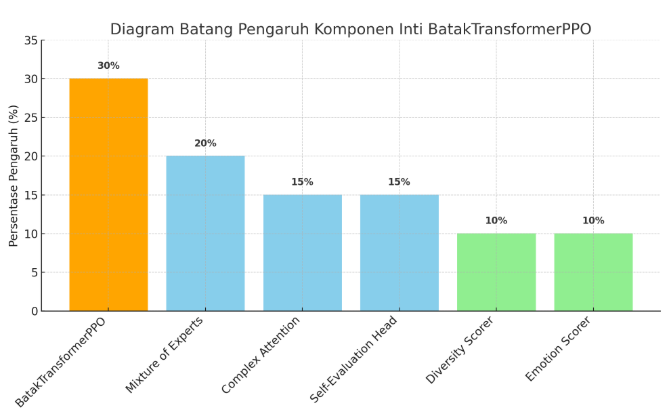
Isi-isi dari MoE adalah Expert Network, yaitu grup model-model kecil yang ditandai dengan *feedforward Neural Networks*, lalu ada juga *Gating Networks,* yaitu layer kecil yang berisi linear dan softmax yang bertugas memilih experts mana yang harus diaktifkan untuk setiap input, agar hemat komputasi. Lalu ada *Weighted Combination* yaitu output(keluaran, hasil) dari expert terpilih dikalikan dengan bobot dari gating, lalu jumlahkan. Maka hasil akhirnya adalah jawaban dari expert terpilih.

Contoh: Apa itu AI.

Maka Experts bertugas untuk mencari jawaban, masing-masing experts mencari “Pengertian AI”, “Fungsi-fungsi AI”, “AI dalam Kantor”, “AI dalam dunia kerja” masuk ke Gating Network untuk menunggu bagaimana experts menentukan jawaban, lalu melalui top\_k, lalu top\_k itu sendiri akan memilih bagian yang relevan, contoh Pengertian dan Fungsi-fungsi nya. experts yang mendapat 2 data lainnya yang dianggap tidak mirip dengan pertanyaan user. Lalu darisana 2 data yang terpisah tadi digabungkan dan jadilah 1 jawaban, dan ini adalah jawaban kenapa pertanyaan teknis mendapatkan banyak data dalam 1 output.

*Analysis - In Indonesian Language*

# Marca - No Human Intervetion



Ini adalah tolak ukur kemampuan penggunaan AI, dimana emotion scorer dan diversity scorer mendapatkan tempat yang paling bawah. Kenapa? Karena diversity dan emotion scorer bukan inti dari model yang hanya *Classification*, yang agar kelihatan bernuansa emosional dan pengecekan terhadap variasi output. Peringkat yang berada di tengah-tengah adalah Self-Evaluation Head dan Complex Attention. Kenapa? Karena jika digabung, 30% Model yang dimiliki menerima Kritik dan Saran dan selalu memerhatikan susunan kalimatnya. Artinya, Kontribusi nya tidak terlalu kelihatan dibanding hasil yang diberikan (sama seperti manusia, ketika menerima kritik, orang yang jenius akan menerima nya dan melatih kemampuan yang dikritik). Lalu, Mixture of Experts berdistribusi hanya sebagai 20%, karena inipun bagian yang tidak kelihatan dimana tiap experts dalam 1 model mencari jawaban, jadi bisa dikatakan tidak terlalu transparan.

BatakTransformerPPO adalah otak nya! Dalam realita, Otak manusia hanya berfungsi 10 dari 100%, tapi jika AI, haruslah 20% karena dirancang untuk lebih dari manusia.

*Discussion - In Indonesian Language*

# Marca - No Human Intervetion

Sistem BatakTransformerPPO merupakan arsitektur model bahasa berbasis Transformer yang diperkuat dengan komponen modular seperti Mixture of Experts (MoE), Complex Attention, dan berbagai kepala evaluasi seperti Self-Evaluation, Emotion Scorer, serta Diversity Scorer. Inti dari model ini adalah kemampuan adaptifnya dalam memilih jalur representasi terbaik untuk setiap input melalui MoE, dan memperluas perhatian konteks melalui mekanisme Complex Attention yang mampu mengintegrasikan memori historis. Semua hasil decoding disalurkan ke berbagai head yang bukan hanya menghasilkan prediksi token, tetapi juga menilai kualitas, keberagaman, dan nuansa emosional dari jawaban yang dibentuk — menciptakan perilaku model yang tidak hanya responsif tapi juga reflektif.

Sistem pelatihan dikendalikan oleh kelas SmartAssistant, yang memproses data dalam format khusus berbasis prompt-response Bataknese serta soal matematika, menggunakan tokenizer Unigram yang dilatih secara lokal. Pembelajaran dilakukan melalui tiga jalur utama: supervised learning berbasis kurikulum (easy → medium → hard), Direct Preference Optimization (DPO) menggunakan ranking buffer, dan Proximal Policy Optimization (PPO) melalui reward internal berbasis self-evaluation. Respons model dievaluasi otomatis — dan jika skornya rendah, dilakukan sampling ulang dengan temperatur lebih tinggi. Setiap masukan pengguna memperbarui memori konteks dan menciptakan pengalaman PPO baru, yang kemudian dikonsolidasikan dalam pelatihan berkala.

Integrasi logika pelatihan reinforcement, evaluasi mandiri, dan pembelajaran berbasis preferensi ini menjadikan marca1b.py bukan sekadar model generatif, melainkan sistem adaptif mandiri yang mampu menilai, memperbaiki, dan mengembangkan perilakunya dari waktu ke waktu. Ia mencerminkan prinsip dasar dari Reinforcement Learning with AI Feedback (RLAIF) dalam versi lokal yang efisien, modular, dan dirancang untuk mendukung bahasa minoritas. Dengan menyimpan log, memperbolehkan umpan balik pengguna, dan menjalankan pembaruan parameter secara berkelanjutan, sistem ini dapat tumbuh menjadi asisten linguistik dan logika yang semakin matang seiring interaksi — menjadikan kode ini sebagai pondasi untuk eksperimen LLM adaptif masa depan.

*Bibliography - In Indonesian Language*

# Marca - No Human Intervetion

Vaswani et al. "Attention Is All You Need." https://arxiv.org/abs/1706.03762

Devlin et al. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." https://arxiv.org/abs/1810.04805

Raffel et al. "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer (T5)." https://arxiv.org/abs/1910.10683

Schulman et al. "Proximal Policy Optimization Algorithms." https://arxiv.org/abs/1707.06347

Christiano et al. "Deep reinforcement learning from human preferences." https://arxiv.org/abs/1706.03741

Rafailov et al. "Direct Preference Optimization: Your Language Model is Secretly a Reward Model." https://arxiv.org/abs/2305.18290

Ouyang et al. "Training language models to follow instructions with human feedback (InstructGPT)." https://arxiv.org/abs/2203.02155

Bai et al. "Training a Helpful and Harmless Assistant with RLHF." https://arxiv.org/abs/2204.05862

Ziegler et al. "Fine-Tuning Language Models from Human Preferences." https://arxiv.org/abs/1909.08593

Kim et al. "Self-Rewarding Language Models." https://arxiv.org/abs/2302.13971

Honovich et al. "True Teacher: Self-supervised evaluation for instruction tuning." https://arxiv.org/abs/2212.09251

Lewis et al. "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks." https://arxiv.org/abs/2005.11401

HuggingFace. "Transformers Documentation." <https://huggingface.co/docs/transformers>

Shazeer, N., Mirhoseini, A., Maziarz, K., Davis, A., Le, Q. V., Hinton, G. E., & Dean, J. (2017). “Outrageously Large Neural Networks: The Sparsely-Gated Mixture-of-Experts Layer.” arXiv preprint arXiv:1701.06538. https://arxiv.org/abs/1701.06538