



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 4 No. 4 (2025) pp: 673-680

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Mengoptimalkan Big Data dan AI dalam Menentukan Calon Mahasiswa Berkualitas sebelum Seleksi Perguruan Tinggi

Gunawan Wibisono, Heri Napitupulu

Universitas Nurtanio Bandung

gunawaib@gmail.com, napitea@gmail.com

Abstrak

Dalam era digital, peningkatan jumlah pendaftar perguruan tinggi menimbulkan tantangan baru dalam proses seleksi mahasiswa. Metode konvensional yang hanya mengandalkan nilai akademik dan ujian tertulis sering kali belum mampu menggambarkan potensi calon mahasiswa secara menyeluruh. Oleh karena itu, penerapan Big Data dan Artificial Intelligence (AI) menjadi solusi inovatif untuk mengoptimalkan proses seleksi calon mahasiswa berkualitas sebelum tahap penerimaan resmi. Melalui pengolahan Big Data, berbagai informasi seperti nilai akademik, aktivitas ekstrakurikuler, rekam jejak digital, portofolio, hingga perilaku belajar daring dapat dianalisis secara komprehensif. Sementara itu, algoritma AI seperti machine learning dan predictive analytics mampu mengidentifikasi pola-pola tertentu yang berkorelasi dengan keberhasilan studi dan kontribusi mahasiswa di masa depan. Sistem ini memungkinkan perguruan tinggi untuk melakukan penilaian berbasis data secara objektif, cepat, dan efisien. Selain itu, integrasi AI juga dapat membantu memberikan rekomendasi personalisasi jalur penerimaan serta meminimalkan potensi bias manusia dalam proses seleksi. Implementasi teknologi ini tidak hanya meningkatkan kualitas input mahasiswa, tetapi juga mendukung terciptanya ekosistem pendidikan tinggi yang adaptif dan berorientasi pada data. Dengan demikian, optimalisasi Big Data dan AI berperan strategis dalam membangun sistem seleksi mahasiswa yang transparan, akurat, serta mampu memprediksi potensi akademik dan non-akademik calon mahasiswa secara lebih menyeluruh.

Kata kunci: Parameter, Big Data, Artificial Intelligence

1. Pendahuluan

Dalam dunia Pendidikan, mahasiswa merupakan objek yang sangat penting untuk dididik dengan harapan mendapatkan hasil didik yang sesuai dan dibutuhkan dalam dunia kerja serta memberi dampak yang positif dan besar terhadap institusi pendidikan untuk digunakan sebagai *feedback* dalam perencanaan pendidikan selanjutnya sehingga sistem dan metode yang dibentuk dapat dikembangkan untuk memenuhi kebutuhan baik kualitas maupun kualitas akan hasil didik berkualitas dalam skala yang lebih besar dan dalam jangka panjang agar dapat mengadopsi setiap perkembangan baik kurikulum, aturan, kebijakan, metode maupun perkembangan teknologi.

Calon mahasiswa sebagai calon peserta didik datang dari berbagai kalangan masyarakat yang berbeda latar belakang budaya, kualitas pendidikan, disiplin, agama, bahasa daerah maupun bahasa asing dan lain sebagainya akan mempengaruhi pelaksanaan pendidikan apabila tidak dilakukan standarisasi atau yang populer dengan istilah matrikulasi, khususnya untuk mengikuti kurikulum pendidikan berbasis STEM (*Science, Technic, Engineering and Mathematic*). Hal yang sering terjadi adalah masalah kemampuan penguasaan dasar ilmu pengetahuan alam, kimia, fisika dan matematika. Untuk membantu mendapatkan mahasiswa yang berkualitas sesuai dengan kemampuan, minat dan bakat atau *passion* secara masif dengan data akurat, sistem *Big Data* dengan proses kecerdasan buatan akan memberi solusi yang terbaik.

Big Data dengan kemampuan menampung data dengan kapasitas yang sangat besar, data tersebut bukanlah data yang terstruktur dan dapat diambil dari manapun sumbernya. Terkait dengan yang akan diproses adalah data calon mahasiswa maka sumber masuk dari *Big Data* dapat dari Disdukcapil, data lulusan sekolah menengah, Lembaga-lembaga kursus atau bimbingan belajar dan lain sebagainya. Dalam memproses data tersebut dapat dilakukan dengan memanfaatkan sistem kecerdasan buatan atau *artificial intelligence* (AI) dengan beberapa metode dan perangkatnya (tools) yang ada. Bagi *Big Data* semua data adalah penting karena data tersebut makin

banyak makin meningkatkan peluang dalam memberi layanan terbaik dan pada akhirnya dapat meningkatkan kinerja suatu unit kerja yang memanfaatkan *Big Data* dengan proses AI.

Profesor John McCarthy (4 September 1927 – 24 Oktober 2011) adalah ilmuwan komputer dan ilmuwan kognitif dari Amerika yang pertama kali memperkenalkan sistem kecerdasan buatan (*artificial intelligence*, AI) di MIT (Massachusetts Institute of Technology) pada tahun 1956 dengan mengembangkan pemrograman logika. Setelah itu kemajuan dibidang kecerdasan berkembang pesat seiring dengan perkembangan teknologi elektronika, perkembangan bahasa pemrograman serta perkembangan Teknik algoritma sampai terciptanya mesin-mesin pemrograman untuk mensimplifikasi sesuatu yang awalnya rumit dibuat sedemikian rupa sehingga menjadi sesuatu yang lebih sederhana dan memudahkan untuk diimplementasikan.

Teori yang digunakan dalam penelitian ini meliputi sistem *Big Data*, *artificial intelligence*, *searching algoritma*, *algoritma AI*, dan *searching engine*.

Big Data. *Database* didasarkan pada data yang teratur untuk memenuhi field, record dan juga file, dapat menampung data dengan kuantitas besar namun masih harus memenuhi jenis data yang sama dan dimasukkan dalam suatu basis data. Besar kecilnya data tergantung dari kemampuan memori yang disediakan oleh sistem computer. Lain halnya dengan Big data yang merupakan pengembangan teknologi dalam sistem pengolahan data. Dalam teknologi Big data pengembangannya sejalan dengan pengembangan pemrosesan atau algoritma yang semakin rumit dan mampu mengenali data lalu menganalisanya untuk selanjutnya mengidentifikasi pengenalan pola. Kualitas data yang digunakan untuk melatih algoritma AI sangat penting. Data yang baik harus relevan, akurat, dan bebas dari bias. Dalam memproses dan menganalisis data besar, dibutuhkan teknologi dan alat yang dapat menangani volume, kecepatan, dan keragaman data. Banyak tersedia alat tersebut bahkan alat tersebut adalah open source, diantara teknologi yang sering digunakan dalam analisis *Big Data* tersebut adalah *Hadoop* dan *NoSQL*.

Hadoop. Hadoop adalah algoritma open-source yang dirancang untuk memproses dan menyimpan *Big Data* dengan mempertimbangkan tingkat efisiensi tinggi. *Hadoop* dirancang dengan model pemrograman *MapReduce* untuk mendistribusikan hasil pemrosesan data ke berbagai simpul dalam *cluster*, sehingga dapat menangani data dengan volume yang besar, cepat dan akurat. *Hadoop* sangat efektif untuk menangani data terstruktur maupun tidak terstruktur, seperti yang sering dijumpai dalam analisis *Big Data*.

NoSQL. Berbeda dengan SQL yang digunakan dalam pengolahan data relasional, NoSQL adalah jenis pengolahan data yang tidak menggunakan model data relasional, yaitu model data yang lebih fleksibel, seperti dokumen, grafik, atau kolom. NoSQL dirancang untuk menangani volume data yang sangat besar dan tidak terstruktur, serta memiliki kecepatan akses yang tinggi. Teknologi ini sering digunakan dalam analisis *Big Data* untuk mendapatkan wawasan yang berharga dari data yang beragam dan tidak teratur. Apabila kebutuhan organisasi terus berkembang dan berinovasi, agar kemampuan kecepatan proses tetap tinggi dalam skala apapun maka dibutuhkan suatu alat yang dapat menangani semua tuntutan maka NoSql akan tampil menawarkan dan melayani semua yang dibutuhkan untuk membangun aplikasi dengan volume data yang besar dan dengan latensi atau waktu respon yang kecil (JSON, 2021)

Artificial intelligence. Profesor John McCarthy (4 September 1927 – 24 Oktober 2011) adalah ilmuwan komputer dan ilmuwan kognitif dari Amerika yang pertama kali memperkenalkan sistem kecerdasan buatan (*artificial intelligence*, AI) di MIT (*Massachusetts Institute of Technology*) pada tahun 1956 dengan mengembangkan pemrograman logika. Setelah itu kemajuan dibidang kecerdasan berkembang pesat seiring dengan perkembangan teknologi elektronika, perkembangan bahasa pemrograman serta perkembangan teknik algoritma sampai terciptanya mesin-mesin pemrograman untuk mensimplifikasi sesuatu yang awalnya rumit dibuat sedemikian rupa sehingga menjadi sesuatu yang lebih sederhana dan memudahkan untuk diimplementasikan.

Perkembangan AI. Perkembangan AI telah berhasil membuat Teknik Informatika menjadi mesin yang siap bergerak dan bekerja sesuai spesifikasinya masing-masing seperti *Machine Learning*, *Deep Learning*, *Neural Network* dan lain lain, yang kesemuanya itu dapat digolongkan atau dikategorikan sebagai kemampuan AI seperti AI dapat mengenali objek berupa citra atau suara yaitu data data digital yang diperoleh dari sensor-sensor yang terhubung ke sistem melalui jaringan syaraf buatan sebagaimana ada dalam organ tubuh manusia yang faktanya adalah suatu sistem komunikasi data yang ditransmisikan baik melalui udara maupun kabel yang dapat berupa kabel tembaga maupun serat optik. Dengan demikian maka AI pada dasarnya meniru sistem kecerdasan manusia berdasarkan ketrampilan kognitif, penalaran, pembelajaran dan sistem evaluasi. Saat ini *Artificial intelligence*

(AI) secara umum didefinisikan sebagai mesin buatan manusia yang dapat meniru kecerdasan manusia dan mampu mengerjakan pekerjaan manusia sesuai dengan algoritma yang memberi instruksi kepada sistem komputer yang mengendalikannya. Proses AI melibatkan beberapa langkah utama, mulai dari pengumpulan data hingga penerapan model AI dalam dunia nyata. Secara umum, AI bekerja dengan cara mengambil data, menganalisisnya untuk mengenali pola, dan menggunakan pola tersebut untuk membuat prediksi atau mengambil keputusan. (Dr. Suyanto, 2021)

Cara Kerja AI. Secara umum AI bekerja dengan mensinergikan data yang terkumpul dengan algoritma yang dibuat untuk mengaktifkan dan mengefektifkan data dengan cara memproses menjadi output. Selanjutnya, AI akan menyerap sejumlah data lalu menganalisisnya untuk mengidentifikasi pengenalan pola. Sementara untuk sistem kerjanya, AI membutuhkan dasar perangkat lunak maupun keras untuk melatih dan menulis algoritma pembelajaran mesin atau *machine learning*. Model belajar AI ini berdasarkan keterampilan kognitif, seperti penalaran, pembelajaran dan koreksi diri.

Algoritma dan Algoritma AI. Kemampuan AI selanjutnya adalah karena adanya kemampuan dari algoritma yaitu berupa kumpulan instruksi digital yang dibuat khusus untuk melakukan kegiatan secara teratur, berurutan dengan mengadopsi logika pembuatnya agar sistem dapat berbuat, bertindak dan mengerjakan semua keinginan programmer atau algoritma, algoritma sering juga dijuluki sebagai sesuatu yang dapat meniru cara berpikir dan bertindak manusia dengan kecepatan tinggi sehingga setiap data yang diterima akan diidentifikasi, dipelajari, diproses, diolah, diprediksi dan dibuat kesimpulan untuk setiap tugas. Dapat disimpulkan bahwa algoritma merupakan hal yang paling mendasar dalam yang memungkinkan sistem dapat menerima (data), berpikir, menyimpulkan dan mengambil keputusan.

Algoritma. Dalam setiap proses komputasi harus ada yang pertama kali dikerjakan sebelum melangkah ke pekerjaan selanjutnya agar semua pekerjaan atau proses berjalan tertib dan sempurna sampai titik urutan terakhir, sehingga pekerjaan berjalan sempurna dan mendapatkan output sebagai solusi. (K, 2024). Sebuah instruksi dapat disebut sebagai algoritma jika memiliki ciri-ciri berikut:

Input: Membutuhkan input (data) untuk diolah.

Proses: Setiap langkah atau instruksi harus jelas dan tidak ambigu.

Output: Menghasilkan keluaran atau solusi dari masalah yang dihadapi.

Keterbatasan: Memiliki urutan operasi yang terbatas dan akan berhenti setelah masalah terpecahkan.

Efektif: Setiap langkah harus dapat dilaksanakan dan menghasilkan keluaran sebagai solusi.

Algoritma AI. Algoritma AI merupakan perkembangan algoritma yang dibuat karena adanya tuntutan pengolahan *Big Data* berupa instruksi atau perintah yang memungkinkan sistem komputer untuk melakukan tugas tugas berat, diarahkan pada meniru karakteristik, cara berpikir dan belajar manusia. Algoritma ini memungkinkan mesin untuk belajar dari tersedianya *Big Data*, membuat prediksi, dan mengidentifikasi pola tanpa perlu pemrograman eksplisit untuk setiap tugas. Contoh penerapan algoritma AI pada berbagai aplikasi seperti pengenalan wajah, meniru gerak manusia pada robot, memodelkan kerja organisasi kedalam analogi sistem jaringan saraf dan sebagainya. Algoritma AI adalah kunci untuk membuat sistem komputer cerdas yang dapat membantu memecahkan masalah dan meningkatkan berbagai aspek kehidupan. Ada beberapa jenis algoritma AI diantaranya adalah pembelajaran mesin (*machine learning*), pembelajaran mendalam (*deep learning machine*), algoritma pencarian (*searching machine*) dan sebagainya. Algoritma AI bekerja dengan menganalisis dan mengeksplorasi data pelatihan untuk menemukan pola dan belajar dalam jangka waktu yang panjang. (tableau, 2021). Algoritma AI membutuhkan data dalam jumlah besar untuk dapat memproses, belajar dan berfungsi secara efektif, diadaptasikan secara matematis untuk membuat prediksi yang akurat. Algoritma sebagaimana manusia harus belajar dari kesalahannya sendiri dengan cara mengambil *feedback* untuk terus meningkatkan kinerjanya. Beberapa penggunaan umum algoritma AI meliputi: *Entry data* dan klasifikasi, analisis lanjutan atau prediktif, Mesin pencari, asisten digital, robotika dan lain-lain.

Searching Engine. *Search engine* tergolong dalam AI yang tanpa disadari paling banyak digunakan, hanya dengan menggunakan kata kunci tertentu, mesin ini dapat menjalankan fungsinya dan memberi informasi dengan baik.

Media yang digunakan umumnya adalah internet atau world wide web (WWW), melakukan indeks kemudian menampilkan hasil pencarian. Hasil pencarian sering disebut dengan SERP (*serach engine result page*).

Hubungan Bigdata dan AI. Bigdata mengacu pada kumpulan data yang masif, kompleks atau tidak terstruktur dan harus diproses dengan kecepatan tinggi untuk menghasilkan informasi bermanfaat sesuai yang diinginkan oleh penerima informasi. Pemrosesan tersebut menggunakan algoritma AI dan mesin pembelajaran untuk menggabungkan dan menganalisis data tersebut agar dapat diidentifikasi pola data tersebut selanjutnya dapat dikembangkan dan ditindaklanjuti. *Big Data* dapat diumpamakan sebagai bahan bakar supaya AI dapat bekerja dan menghasilkan produk sesuai dengan keinginan. Hal ini diperlukan agar hasil proses tersebut dapat digunakan dalam pengambilan keputusan yang cepat agar dapat meningkatkan kinerja. *Big Data* menyediakan data yang massif, data tersebut dianalisis oleh AI untuk menentukan pola agar dapat memprediksi serta mengambil keputusan yang tepat dan cerdas. Dengan kata lain *Big Data* menyediakan volume data masif dan bervariasi yang dibutuhkan AI untuk menjalankan algoritmanya, mendapatkan wawasan dari data tersebut, menganalisa secara optimal sehingga dapat menghasilkan keputusan secara tepat sasaran dalam waktu yang sangat singkat atau *real time*. *Big Data* dan AI bekerjasama secara berkesinambungan dalam melaksanakan tugasnya masing-masing antara lain seperti berikut :

Fungsi. *Big Data* bertugas dan berfungsi mengumpulkan dan menyimpan data sementara AI melaksanakan pemodelan sesuai perintah algoritma sesuai data yang teridentifikasi serta menemukan data tersembunyi, mendeteksi anomali dan membangun model untuk tugas-tugas seperti prediksi, klasifikasi, dan segmentasi.

Sinergi. *Big Data* menyediakan data sementara AI melaksanakan analisis prediktif yang berwawasan kedepan selain menganalisis apa yang telah terjadi, sinergi ini akan bekerja secara saling memenuhi akan tuntutan tugas yang dibebankan kepada masing-masing.

Berkelanjutan. Tugas-tugas kompleks yang dikerjakan oleh AI dalam menganalisis data dilaksanakan berkelanjutan secara otomatis tanpa adanya campur tangan operator atau perintah tertentu yang menyela di tengah proses yang dikerjakan oleh operator atau prmbuat program

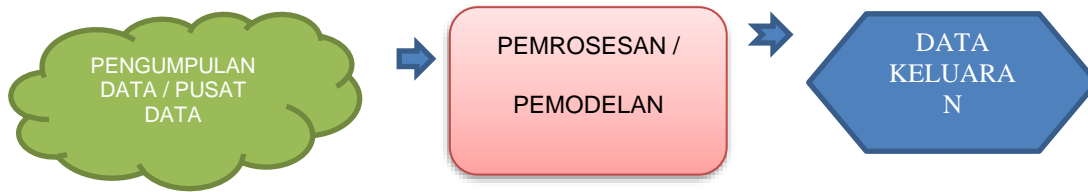
2. Metode Penelitian

Penelitian ini menggunakan metode analisis, sintesis literatur dan analogi dari organisasi sistem berjalan menjadi model keluaran. Langkah pertama adalah mengidentifikasi sumber-sumber literatur yang relevan melalui pencarian online yang meliputi buku, jurnal ilmiah, repositori universitas, dan situs-situs terpercaya lainnya. Setelah sumber-sumber yang tepat telah diidentifikasi, dilakukan analisis terhadap konten yang terkait dengan Manajemen bigdata, sebagai wacana konsep memvisualisasikan konsep. Dalam Konteks Sistem Informasi Manajemen, struktur organisasi dan tugas pokok organisasi menjadi acuan dalam memvisualisasikan kerja organisasi dianalogikan menjadi model sistem yang akan dibangun. Beberapa hal penting yang harus dilakukan dalam proses AI:

Pengumpulan dan Persiapan Data. Langkah pertama adalah mengumpulkan data yang relevan dan membersihkannya (*data preprocessing*) agar siap digunakan oleh model AI. Data tersebut diperoleh dari berbagai sumber dan sifatnya massif, tidak terikat oleh field sehingga data apa saja dapat diambil dengan kriteria tertentu yang mendekati keinginan algoritma AI, sebagai upaya menyediakan data yang akan dapat digunakan pada proses AI.

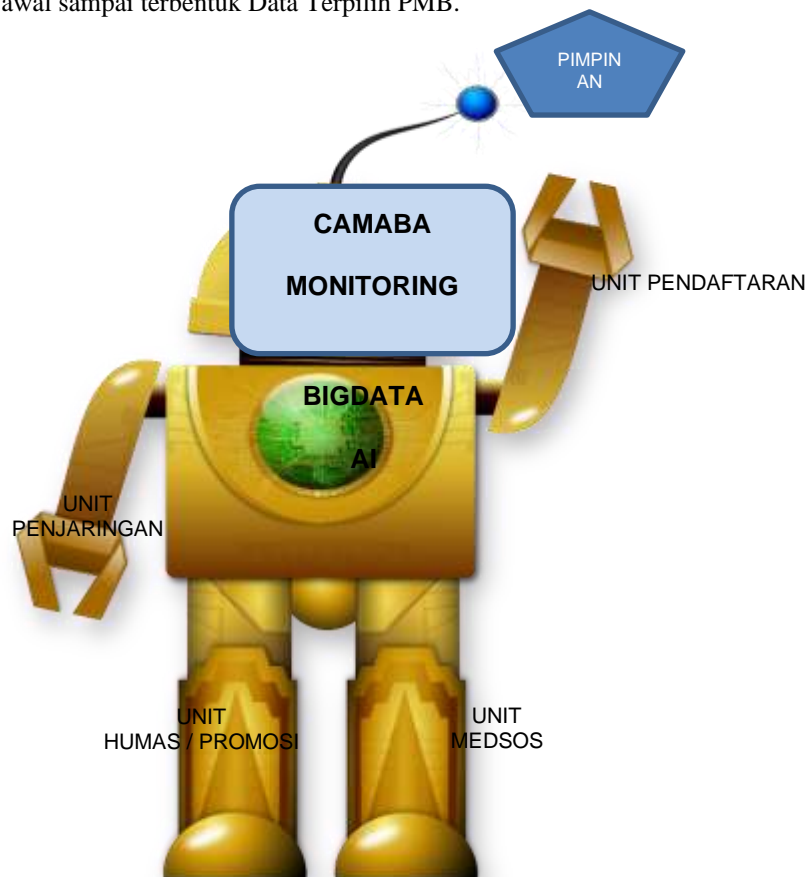
Pemodelan AI. Model AI, seperti halnya jaringan saraf tiruan (neural network) sebagai contoh, dikembangkan menggunakan algoritma AI. Algoritma ini dilatih pada data yang tersedia untuk dipelajari dan dievaluasi untuk mengenali pola sehingga dapat digunakan untuk membuat prediksi.

Data Keluaran. Model AI yang telah diuji dan dievaluasi kemudian dikumpulkan dalam suatu penyimpanan yang selanjutnya akan mendapat status tertentu sebagai data terpilih, contohnya adalah data calon mahasiswa awal dan akan dimanfaatkan untuk intersepsi lebih lanjut terhadap data yang memiliki karakteristik sejenis.



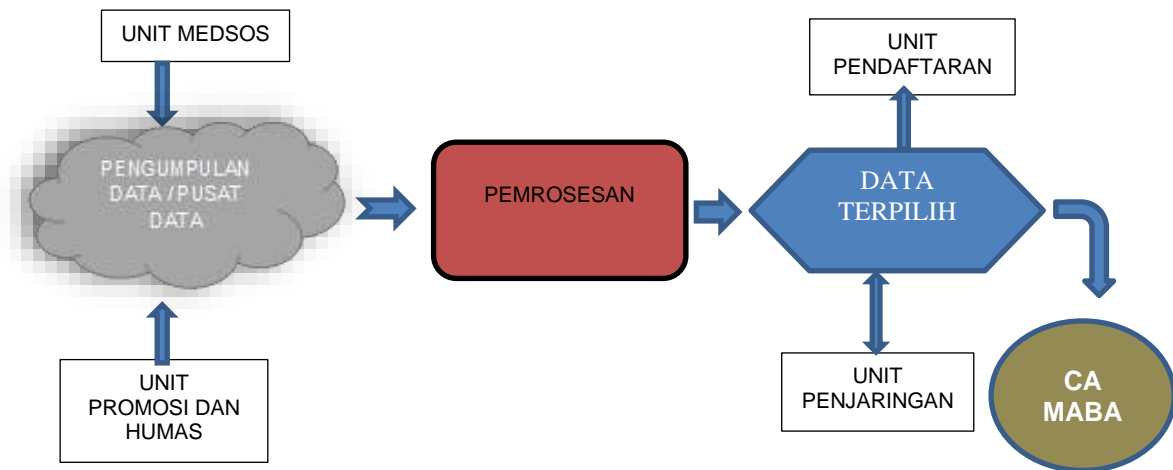
Gambar 1. Implementasi Metode Penelitian

Konsep Hardware. Memvisualisasikan atau menganalogikan struktur organisasi yang terkait dengan Penerimaan Mahasiswa Baru (PMB) sebagai suatu struktur nyata/hidup yang terhubung dengan pusat kendali/otak yang memiliki memori untuk menyimpan data mentah kemudian dilakukan proses atau pengolahan data yang dilaksanakan oleh algoritma AI, terhubung dengan bagian-bagian unit kerja yang berfungsi mencari, menseleksi, mengolah data tahap awal sampai terbentuk Data Terpilih PMB.



Gambar 2. Model Konsep Hardware

Unit-unit kerja melakukan interaksi dan memberi masukan data dengan sistem AI untuk membuat sistem kerja yang terkendali sehingga dapat memberikan keluaran berupa data terpilih sebagai dasar atau modal dalam melaksanakan kerja PMB secara terstruktur dan terarah.



Gambar 3. Model Konsep Implementasi

3. Hasil dan Diskusi

Pusat Data. Pusat data yang memfasilitasi *Big Data* mendapat masukan dari berbagai sumber data baik secara online maupun offline. Data tersebut merupakan data kotor dalam hal ini dari semua lulusan SMA/SMK, calon mahasiswa yang gagal masuk Perguruan Tinggi Negeri (PTN) atau karyawan yang masih mungkin untuk melanjutkan kuliah setingkat S1/D4 ataupun Pasca Sarjana dan lain sebagainya.

Unit Media Sosial. Media social merupakan sumber data potensial yang dapat dimanfaatkan oleh Tim PMB untuk diproses oleh AI. Media sosial merupakan arena promosi dari Tim PMB Unnur, siapapun yang mengakses atau merespons materi promosi PMB Unnur harus dapat terjaring masuk dengan melabelkan data pribadi seperti nama, alamat, nomer hp dan sebagainya

Unit Promosi dan Humas. Merupakan unit operasional melakukan promosi dan penyebaran informasi kepada siapapun yang berpotensi menjadi calon mahasiswa, mendapatkan data sebanyak mungkin dengan data yang dapat digunakan sebagai modal pencarian selanjutnya diolah oleh AI untuk mendapatkan data yang lebih lengkap lagi dengan memanfaatkan search machine.

Unit Penjaringan. Merupakan unit yang berfungsi secara interaktif mengamati data yang disajikan dari hasil pemrosesan untuk selanjutnya menentukan kriteria, memilah atau menseleksi data yang akan dinominasikan sebagai calon mahasiswa baru serta berupaya melakukan penggiringan agar semua yang sudah masuk wadah data terpilih untuk bias mendaftar sebagai mahasiswa baru.

Unit Pendaftaran. Adalah unit yang melaksanakan penerimaan calon mahasiswa baru dengan kegiatan menampung semua data terpilih untuk crosscheck dengan yang melakukan registrasi sebagai bahan kaji ulang atau sebagai feedback. Unit Pendaftaran memfasilitasi dan memproses semua calon mahasiswa baru yang akan melakukan registrasi.

Konsep Software. Konsep software merupakan otak yang bekerja di balik layar dan mempunyai peran yang sangat besar sampai terciptanya sistem cerdas. Dari konsep software ini suatu karya sistem cerdas dapat menunjukkan performanya dan dari sini pula sistem cerdas dapat ditingkatkan kemampuannya.

Pengumpulan Data. Data yang akan diproses adalah data dari berbagai sumber baik diperoleh secara online maupun offline. Data dari online diperoleh dari platform medsos, saat membuka platform medsos maka AI akan melihat banyak data sebagai informasi baru, kemudian AI yang berfungsi sebagai mesin pencari (search engine) dengan menggunakan kata kunci yang sudah diberikan akan mencari dan memilih data data yang diperlukan. Bahkan AI dapat memperoleh data wajah dengan menggunakan fitur Face ID. Dengan perputaran dan peredaran data di web yang semakin massif maka data yang terkumpul pun akan sangat banyak. Pengumpulan data ini juga akan semakin banyak dan akurat dengan adanya masukan dari Tim Humas yang akan memberi referensi pencarian.

Pemrosesan. Dari data massif yang sudah terkumpul selanjutnya adalah memproses data dengan mengedepankan peran algoritma AI untuk mencari pola, kemudian dengan menggunakan informasi tersebut digunakan untuk memprediksi atau mengambil keputusan atau tindakan. Algoritma AI dapat difungsikan sebagai suatu mesin seperti mesin pembelajaran, pembelajaran mendalam, mesin pencarian dan sebagainya. Disinilah algoritma AI menunjukkan peran sebagai kunci dalam sistem cerdas yang dapat membantu memecahkan masalah sehingga dapat membantu meningkatkan kinerja.

Data Keluaran. Dari hasil pemrosesan data selanjutnya ditampung dalam suatu database, selanjutnya data ini digunakan oleh Unit Penjaringan untuk ditentukan susunan prioritas calon mahasiswa agar efektif dan efisien dalam pelaksanaan intersepsi lapangan.

Konsep Model. Membuat suatu model yang dibangun harus menggambarkan sistem nyata yang bekerja seperti halnya manusia atau robot yang dikendalikan oleh software, terhubung kesemua bagian seolah memiliki syaraf dan setiap bagiannya sebagai client juga memiliki sensor, pengolah, dan memori lokal. Dalam model AI, pada bahasan *Machine Learning* (ML) harus menjadi perhatian khusus, karena harus merancang dan membangun algoritma yang akan melatih, menguji dan menyetel hingga sempurna. Hal tersebut harus demikian karena proyek yang dibuat ini hasilnya harus meniru seperti layaknya manusia dan memiliki karakteristik seperti apa yang diinginkan oleh pembuatnya. Jangan abaikan *Questing Answer* (QA) pada setiap tahapnya. Tahapan yang harus ditempuh dalam

Bertanya Dengan Tepat. Tahap pertama ini berarti menanyakan pertanyaan yang tepat yaitu melakukan identifikasi masalah dengan mencari dan mendapatkan data calon mahasiswa yang keberadaannya ada dimana mana dengan memanfaatkan machine searching yang ada di media sosial. Menanyakan berulang terhadap data yang sudah terkumpul sekaligus mengevaluasi dan memverifikasi data agar data menjadi bersih dan terpilih. Tahap ini adalah tahap terpenting karena data yang terpilih merupakan modal dasar pekerjaan tim untuk ditindak lanjuti. Data tersebut berupa data calon mahasiswa dengan data seperti usia, pendidikan terakhir yang masuk kriteria, asal sekolah, tahun lulus SMA/K, nomor handphone dan lain-lain data terkait dengan kriteria calon mahasiswa yang diharapkan

Penyiapan Data. Untuk melakukan penyiapan data yaitu dengan mengambil langkah berikutnya dengan menentukan kriteria yang tegas sehingga system AI dapat menentukan data terpilih yang memenuhi kriteria dan system AI dapat mengurutkan hasil berdasarkan kualitas data terhadap kemungkinan diperolehnya calon mahasiswa sesuai yang dikehendaki oleh Unnur dengan nilai atau kualitas untuk ditempatkan dan diurutkan dari yang teratas sampai terendah berdasarkan kemungkinan didapat (secara statistik). Data yang terkumpul ini akan mencerminkan kemungkinan perolehan calon mahasiswa dengan memanfaatkan *Machine Learning* (ML). Pengumpulan data ini dilakukan berkali kali dan harus dilaksanakan secara teratur agar data yang diproses oleh akan menghasilkan data yang berkualitas tinggi. Semakin banyak data semakin baik, akan menghasilkan data yang relevan dan berkualitas tinggi.

Menentukan Algoritma(Algoritma Data Science Education Center, 2022). Algoritma AI yang akan digunakan dalam suatu proyek harus memperhatikan atau mempertimbangkan banyak factor. Algoritma AI ini adalah *machine learning* yang harus bekerja efisien dan efektif dalam membaca dan memprediksi data. Dalam penerapan Algoritma *machine learning* yang merupakan pengembangan dari *data science* dan AI harus dipelajari secara mendalam dan harus disesuaikan dengan materi yang ditangani. Berikut adalah teknik yang ditempuh:

Apabila permasalahan yang sangat beragam sudah dapat diidentifikasi seperti faktor klasifikasi, regresi, deteksi *anomaly*, *clustering* dan sebagainya maka algoritma yang tidak diperlukan dapat dieliminasi karena hal ini dapat mengefisienkan waktu dan pemetaan.

Mempertimbangkan ukuran *training set*, dalam pengumpulan data dapat diidentifikasi seberapa besar data set yang akan diproses. Dalam hal calon mahasiswa baru, data set kotor yang dikumpulkan diperkirakan tidak terlalu banyak namun dengan fitur yang cukup lengkap maka algoritma dengan *bias* dan *varian* yang rendah seperti *Linear SVM* dan *Naïve Bayes* dapat dipilih.

Dalam proses kerja algoritma, setiap algoritma memiliki kecepatan waktu atau *process training* yang berbeda beda. Pilihan model algoritma seperti model *Naïve Bayes*, *Linear Logistic*, dan *Linear Regression* mudah diimplementasikan dan lebih cepat serta tidak memerlukan parameter tertentu.

Dalam memilih algoritma *machine learning* harus dipastikan tingkat akurasi. Dalam penerimaan calon mahasiswa baru, penentuan tingkat akurasi tidaklah terlalu tinggi, namun dalam teknik informatika yang membahas algoritma, untuk menentukan tingkat akurasi harus membandingkan dengan model lain atau mempelajari algoritma lain yang telah digunakan pada aplikasi AI lainnya agar efektifitas dan selektifitas dalam memilih data dari data terkumpul menghasilkan hasil yang proporsional.

Dalam memilih algoritma, linearitas data juga merupakan hal penting untuk diperhatikan. Adakalanya algoritma yang tidak menggunakan prinsip linearitas tidak terlihat adanya masalah yang berarti, namun di sebagian model lainnya sebagai kasus memang akan berdampak pada tingkat penurunan akurasi. Ada beberapa algoritma yang menerapkan system linearitas antara lain *Linear Regression*, *Logistic Regression*, hingga *Support Vector Machine (SVM)*.

Dalam proses penerimaan mahasiswa baru, perlu dikenali fitur yang dimiliki. Ada kemungkinan data memiliki banyak fitur namun belum tentu bisa digunakan atau dengan kata lain tidak relevan. Pada data yang berupa tekstual, banyak fitur yang harus dikomparasikan dengan sejumlah data lainnya agar tidak menghambat proses penggunaan algoritma *machine learning*. Proses PMB memiliki fitur besar dan pengamatan data sedikit sehingga akan lebih tepat bila menggunakan model SVM karena model algoritma *machine learning SVM (Support Vector Machine)* digunakan untuk klasifikasi dan regresi, dengan tujuan mencari batas pemisah terbaik.

4. Kesimpulan

Dalam membangun sistem dengan memanfaatkan teknologi AI dan *Big Data* perlu dicermati sistem yang akan dibangun karena dengan mencermati sistem yang akan dibangun maka akan diperoleh efisiensi dari penggunaan software AI. Data yang diperlukan dalam Penerimaan Mahasiswa Baru memang harus diproses dari data yang besar atau bahkan sangat besar dan tidak relasional maka teknologi *Big Data* yang digunakan ada beberapa pilihan. Untuk lebih mengefisienkan penggunaan anggaran maka dapat digunakan teknologi *Big Data* yang *open source*.

Referensi

1. Shah, A., & Patil, M. K. (2020). Comparative Study of SQL and NoSQL Databases: Concepts, Architecture, and Applications. *Journal of Computer Science and Technology*, 25(4), 112-123.
2. Gupta, R., & Sharma, S. (2021). Choosing the Right Database: A Guide to SQL vs NoSQL for Modern Applications. *Journal of Software Engineering and Technology*, 18(3), 87-99.
3. Khan, A. L., & Ali, F. (2022). SQL vs NoSQL: An In-Depth Review of Performance, Scalability, and Flexibility. *International Journal of Database Management*, 34(2), 45-59.
4. Patel, S., & Mishra, K. (2023). Adopting NoSQL in Big Data Analytics: Challenges and Future Directions. *Journal of Big Data Research*, 9(1), 52-66.
5. Kumar, S., & Verma, P. (2024). Database Technologies for the Future: A Focus on SQL and NoSQL in Cloud Environments. *Journal of Cloud Computing and Data Storage*, 5(2), 34-47.
6. Dina, A. R. (2025). Pemanfaatan Big Data untuk Kesuksesan Mahasiswa dan Pertumbuhan Institusi. *Jurnal Mentari*, 3(2). DOI: 10.33050/mentari.v3i2.746
7. Xie, Q., Qian, & Wang. (2021). Implementasi Big Data Analytical untuk Perguruan Tinggi Menggunakan Machine Learning. *Jurnal JIFORTY*, 2(1), 77-88.
8. Nadya, N. S. (2024). Analisis Tren Pendaftaran Siswa Menggunakan Big Data di Yayasan Pendidikan Raksana Medan. *Jurnal FKPT*.
9. Hendikawati, P. (2025). Pembelajaran Big Data di Perguruan Tinggi. *Jurnal STKIP PGRI Tulungagung*.
10. Qomariyah, F. (2024). Analisis Big Data dalam Peminatan Mahasiswa Baru. *Jurnal IAIN Madura*.
11. Rachmatika, P. A. (2025). Analisis Penerimaan Mahasiswa Perguruan Tinggi dengan Model UTAUT. *Journal of Engineering, University of Lampung*.
12. Vinay et al. (2015). Pengenalan Wajah Menggunakan Big Data. [Jurnal terkait]
13. Jamshidi, M. (2017). Kerangka Sistem Otonomi Berbasis Big Data. [Jurnal terkait]
14. Sudirman et al. (2024). Penggunaan AI untuk Prediksi Kinerja Akademik di Perguruan Tinggi. [Jurnal terkait]
15. Utami, L. P. (2023). Machine Learning dalam Seleksi Mahasiswa Baru. *Jurnal Teknologi dan Pendidikan*.
16. Harahap, R. (2024). Dasbor Berbasis AI untuk Monitoring Mahasiswa. *Jurnal Sistem Informasi*.
17. Pratama, I. (2025). Optimalisasi Sumber Daya Perguruan Tinggi dengan Big Data. *Jurnal Manajemen Pendidikan*.
18. Sari, D. R. (2024). Prediksi Akademik Mahasiswa dengan Algoritma Random Forest. *Jurnal Komputasi*.
19. Yulianto, H. (2023). AI dan Big Data untuk Pengambilan Keputusan di Pendidikan Tinggi. *Jurnal Informatika*.
20. Dharmawan, A. (2024). Implementasi Data Mining dalam Seleksi Calon Mahasiswa. *Jurnal Data Science*.