

## Model Comparison Table

Question	Model 1 (Mistral-Nemo-Instruct-2407) Response	Model 2 (Gemma-2.2b) Response	Model 3 (Llama-3.2B) Response	ChatGPT Response	Ground Truth (From Dataset)	Best Model (Clarity)	Best Model (Accuracy)	Best Model (Speed)	Best Overall Model
Can you explain this document?	<p>This document appears to be a tab-separated values (TSV) file containing employee data, with each line representing an employee record and fields separated by tabs. Here's a breakdown of the columns:</p> <p>Employee_Name: The name of the employee.</p> <p>EmpID: A unique identifier for the employee.</p> <p>MarriedID: Not clear what this field represents.</p> <p>MaritalStatusID: An ID representing the employee's marital status.</p> <p>GenderID: An ID representing the employee's gender.</p> <p>EmpStatusID: An ID representing the employee's employment status.</p> <p>DeptID: The department ID where the employee works.</p> <p>PerfScoreID: An ID representing the employee's performance score.</p> <p>FromDiversityJobFairID: An ID indicating whether the employee was hired through a diversity job fair (0 = No, 1 = Yes).</p> <p>Salary: The employee's salary.</p> <p>TermID: Not clear what this field represents.</p> <p>PositionID: An ID representing the employee's position.</p> <p>Position: The employee's job title.</p> <p>State: The state where the employee works.</p> <p>Zip: The ZIP code of the location where the employee works.</p> <p>DOB: The employee's date of birth.</p> <p>Sex: The employee's biological sex (M = Male, F = Female).</p> <p>MaritalDesc: A description of the employee's marital status (e.g., Single, Married, Divorced).</p> <p>CitizenDesc: A description of the</p>	<p>The document appears to be a table of employee data from an organization. Here's what we can infer from the provided example:</p> <p>Data Fields: The document lists many different fields, including demographic information (like name, gender, marital status), job details (position, salary, department), and performance-related data (performance score, engagement survey scores).</p> <p>Employee Records: Each row likely represents an individual employee with their unique identifier (EmpID) and other personal/professional details.</p> <p>Possible Use Cases: This data could be used for various purposes such as:</p> <p>HR Analysis: Analyzing trends in demographics, salaries, promotions, performance evaluations.</p> <p>Employee Management: Tracking attendance, employee engagement, and</p>	<p>Based on the information provided, it appears that the "document" is a spreadsheet or database file containing employee data for a company. The rows in the file represent individual employees, with various attributes such as name, ID number, marital status, department, salary, and employment history.</p> <p>The columns seem to be organized into several categories:</p> <p>Basic information: Employee Name, ID Number, Date of Birth, Sex, Marital Status, Department.</p> <p>Salary and benefits: Salary, Performance Score, Engagement Survey, Emp Satisfaction.</p> <p>Employment details: Date of Hire, Date of Termination, Reason for Termination, Employment Status, Manager's Name and ID.</p> <p>Work-related information: Job Title, State, Zip Code.</p> <p>The file also includes some additional columns that seem to be related to diversity metrics:</p> <p>Diversity metrics: Hispanic/Latino Race/Desc</p>	<p>This document appears to be an HR dataset containing information about 311 employees. It includes various attributes such as:</p> <p>Personal Information: Employee name, gender, marital status, race, citizenship status, date of birth.</p> <p>Employment Details: Employee ID, department, position, employment status, salary, hire date, termination date (if applicable), reason for termination.</p> <p>Performance &amp; Attendance: Performance score, engagement survey results, employee satisfaction, last performance review date, absences, days late in the last 30 days.</p> <p>Management &amp;</p>	<p>The csv file is data based off of Employee data that has factors like Personal Information, Work details &amp; perofrmance, Diversity details, and their assoicated manager.</p>	Mistral-Nemo	Mistral-Nemo	ChatGPT	Mistral-Nemo

Question	Model 1 (Mistral-Nemo-Instruct-2407) Response	Model 2 (Gemma-2.2b) Response	Model 3 (Llama-3.2B) Response	ChatGPT Response	Ground Truth (From Dataset)	Best Model (Clarity)	Best Model (Accuracy)	Best Model (Speed)	Best Overall Model
	<p>employee's citizenship status (e.g., US Citizen).</p> <p>HispanicLatino: Indicates whether the employee is Hispanic or Latino (No = Not Hispanic/Latino, 1 = Hispanic/Latino).</p> <p>RaceDesc: The employee's race/ethnicity.</p> <p>DateofHire: The date when the employee was hired.</p> <p>DateofTermination: The date when the employee was terminated (blank or "N/A-StillEmployed" if still employed).</p> <p>TermReason: The reason for termination, if applicable.</p> <p>EmploymentStatus: The current employment status (e.g., Active, Terminated).</p> <p>Department: The name of the department where the employee works.</p> <p>ManagerName: The name of the employee's manager or supervisor.</p> <p>ManagerID: A unique identifier for the employee's manager or supervisor.</p> <p>RecruitmentSource: The source through which the employee was recruited (e.g., LinkedIn, Indeed).</p> <p>PerformanceScore: The employee's performance score (e.g., Exceeds, Fully Meets, Needs Improvement).</p> <p>EngagementSurvey: Not clear what this field represents.</p> <p>EmpSatisfaction: Not clear what this field represents.</p> <p>SpecialProjectsCount: The number of special projects the employee has participated in.</p> <p>LastPerformanceReview_Date: The date of the last performance review.</p> <p>DaysLateLast30: The number of days late in the last 30 days (for attendance tracking).</p> <p>Absences: Not clear what this field represents.</p>	<p>performance improvements.</p> <p>Recruitment &amp; Talent Acquisition: Comparing applicant profiles against available job positions and evaluating candidate performance.</p> <p>Additional Information Needed</p> <p>To fully understand the document, you need more context:</p> <p>Data Source: Where did this data come from? Is it a database, spreadsheet, or something else?</p> <p>Purpose of the Data: What is the primary purpose of collecting and using these employee records?</p> <p>Please provide additional information if you have it.</p>		<p>Recruitment: Manager name and ID, recruitment source (e.g., LinkedIn, Indeed), special projects count.</p> <p>Diversity &amp; Inclusion: Indicators for diversity hiring and Hispanic/Latino identification.</p>					

Question	Model 1 (Mistral-Nemo-Instruct-2407) Response	Model 2 (Gemma-2.2b) Response	Model 3 (Llama-3.2B) Response	ChatGPT Response	Ground Truth (From Dataset)	Best Model (Clarity)	Best Model (Accuracy)	Best Model (Speed)	Best Overall Model
What are the names of employees whose names start with 'E'?	Based on the provided citations, the employees whose names start with 'E' are: "Elijian Clukey" "Karthikeyan Ait Sidi" (although his first name starts with 'K', his last name starts with 'E') "Spencer Cole" "Michael Corleone" "Angela Erilus" "Miguel Estremera" "April Evensen" "Susan Exantus"	Based on the provided citations, the names of employees whose names begin with "E" are: Erilus, Angela Estremera, Miguel Evensen, April Exantus, Susan Let me know if you have any other questions.	Based on the provided citations, the employee names that start with 'E' are: "Erilus, Angela" "Estremera, Miguel" "Evensen, April"	Here are the names of employees whose names start with 'E': Eaton, Marianne Engdahl, Jean England, Rex Erilus, Angela Estremera, Miguel Evensen, April Exantus, Susan	Eaton, Marianne Engdahl, Jean England, Rex Erilus, Angela Estremera, Miguel Evensen, April Exantus, Susan	ChatGPT	ChatGPT	ChatGPT	ChatGPT
What is the highest salary, and who earns it?	Based on the provided citations, the employee with the highest salary is "Shana Maurice" with a salary of \$54132.	Based on the provided data, the highest salary listed is \$104,437. This belongs to Ait Sidi, Karthikeyan in his position as a Sr. DBA.	Based on the provided citations, there is no clear indication of the employee's current salary status or any changes. However, one citation shows that "Langton, Enrico" has a Salary of \$46120, which is likely to be an older salary since this is listed first in the list. However, there are more recent salaries mentioned. The highest salary mentioned in the citations is \$81584 for "Le, Binh", who earns it as a Senior BI Developer.	The highest salary is \$250,000, and it is earned by Janet King.	The highest salary is 250,000. The person is Adinolfi, Wilson K	Llama-3	ChatGPT	ChatGPT	ChatGPT
Name one employee with a salary above \$100,000.	Based on the provided citations, "Le, Binh" has a salary of 81584, which is equivalent to approximately 81,584(roundedfrom81583.68),not exceeding100,000.	Based on the provided data, Ait Sidi, Karthikeyan has a salary of \$104,437.	Based on the provided citations, I found one employee with a salary above 100,000. Theemployeeis"Gonzalez,Juan" (10300),whohasasalaryof68898 as a Production Technician II. However, another citation shows Gonzalez's Salary to be \$59892. Additionally, there are no other salaries above \$100,000 mentioned in the citations provided.	One employee with a salary above \$100,000 is Ait Sidi, Karthikeyan, who earns \$104,437.	Adinolfi, Wilson K Ait Sidi, Karthikeyan Akinkuolie, Sarah Alagbe,Trina Anderson, Carol Anderson, Linda Andreola, Colby Athwal, Sam Bachiochi, Linda Bacong, Alejandro Baczenski, Rachael Barbara, Thomas Barbossa,	Llama-3	ChatGPT & Llama-3	ChatGPT	Llama-3

Question	Model 1 (Mistral-Nemo-Instruct-2407) Response	Model 2 (Gemma-2.2b) Response	Model 3 (Llama-3.2B) Response	ChatGPT Response	Ground Truth (From Dataset)	Best Model (Clarity)	Best Model (Accuracy)	Best Model (Speed)	Best Overall Model
					Hector Barone, Francesco A Barton, Nader Bates, Norman Beak, Kimberly Beatrice, Courtney Becker, Renee Becker, Scott Bernstein, Sean Biden, Lowan M Billis, Helen Blount, Dianna Bondwell, Betsy				
Which department has the highest average performance score?	Based on the provided citations, the employees' departments and their average performance scores are as follows: Production: Exceeds: 4 (average) Fully Meets: 3.675 (average of 3.02, 3.60, and 4.13) IT/IS: Fully Meets: 4.96 Software Engineering: Exceeds: 4.50 Needs Improvement: 2.30 Based on the limited data available, if considering only the 'Exceeds' and 'Fully Meets' categories, IT/IS has the highest average performance score with an average of 4.96.	The provided data doesn't contain enough information to determine which department has the highest average performance score. Here's why: Performance Score Data: The dataset only shows some individual employee performance scores, not an overall department average. We would need a table that lists all employees with their departments and average PerformanceScores for each department. Missing Information: The citations include the 'PerfScoreID', but they don't provide the actual	Based on the provided citations, I was unable to find any information about specific departments or their corresponding performance scores. However, we can infer that there is at least one employee in Production with a high performance score (4.60 and 5.00). But we don't have enough data to determine which department has the highest average performance score.	The department with the highest average performance score is Software Engineering.	Software Engineering	ChatGPT	Mistral-Nemo	ChatGPT	ChatGPT

Question	Model 1 (Mistral-Nemo-Instruct-2407) Response	Model 2 (Gemma-2.2b) Response	Model 3 (Llama-3.2B) Response	ChatGPT Response	Ground Truth (From Dataset)	Best Model (Clarity)	Best Model (Accuracy)	Best Model (Speed)	Best Overall Model
		data points associated with it.							

Reflection one: Reflect One: The advantages of using Local LLM's is that your able to use it offline. Since the LLM runs locally on your computer there is no need for internet. Additionally, you are able to customize the model more. Since the Model is your own you can download, you can train/fine tune it with little restrictions. Most LLM's on Local LLM are downloadable, so it's possible. A major disadvatnage is the Space. To download LLM's locally they take up a decent amount of Space, and processing power. This is one disadvantage of you don't have enough space for it. There is also a learning curve for it. If you want to improve the models, you'll need to learn ML techniques, fine tuning methods and other ways to improve the LLM's. Having a local LLM also means to go through a learning curve.

Reflection two: The Local LLM's do worse compared to LLM's online like ChatGPT. ChatGPT is faster and more accurate then the other Local Models. Models like Mistral-Nemo is slow when it comes to understanding my prompts and output an answer, while Gemma is quick but is inaccurate. There are limitations of RAG locally because of how things are stored locally. ChatGPT, Gemini, and other LLM's are run through optimized cloud services, allowing for quick and accurate outputs, but Local LLM's are run on a regular computer which may not be optimized to support LLM. LLM require alot of computing power, a laptop may not have the optimized settings or performances to support LLM's RAG. Locally, the best model was Mistral-Nemo. Although it was very slow, it produced detailed outputs every time with close accuracy, and even beat out chatGPT for the first question. I think it's best Local one because it's trained for accuracy, and not speed.