Interview Evaluation

|  |  |  |
| --- | --- | --- |
| **URL of the interview recording :** | | |
| **Interviewed On** |  |  |
| **Interviewed by (FULL name)** |  |
| **Candidate Name**  **Job Title** | Data Scientist |
| **Candidate status** | | |
| **Summary notes from interviewer (minimum 800 words):**  Narender has completed Bachelor of Technology (B.Tech) in Mechanical Engineering. Has over ~ 4 years in data engineering, data analysis and data scientist. Worked on sales forecasting, propensity models and Market Basket Analysis models.  Worked on the time series forecasting models, regression, clustering and classification algorithms.  Coding experience : : Python, SQL (MySQL, Snowflake, DuckDB), R (Basic)  Visualization: Power BI  **Having good understanding on base Statistical concepts, machine learning algorithms, coding experience in python/sql, hence selected for Next round.**  His performance is rated below for documentation purpose.  **Communication Skills – /10**  **Logical Thinking – /10**  **Confidence – /10**  **Clarity of thought – /10**  **Attitude – /10** | | |

**Ratings along with comments:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sl. #** | **Skill** | **Rating 0-10**  **(10 is best)** | **Subjective feedback with examples** |
| 1 | What is the formula for Variance and Standard Deviation | 8 | Explained |
| 2 | Methods to apply for treating missing values | 6 | Mean imputation method explained |
| 3 | Explain ANOVA, P - Value and Central limit theorem. How to interpret the ANOVA table | 5 | Partially explained |
| 4 | Define the Null Vs Alternate Hypothesis,  Difference between Type 1 & Type II error | 6 | Partially explained |
| 5 | Explain the assumptions of Multiple Linear Regression model | 5 | Explained only two assumptions |
| 6 | What is the difference between R square and adjusted R square  When to use R2 vs MAPE, RMSE etc | 5 | Partially explained |
| 7 | How can we check and remove Multi-collinearity from data?  What is the threshold value for VIF? | 8 | Explained the concepts, method to handle with the dataset |
| 8 | Difference between Random forest Vs logistic regression models? Evaluation metrics used for classification algorithms. | 8 | Explained, Precision and Recall, F1 score |
| 9 | Time Series Forecasting | 8 | Experienced in time series forecasting models. |
| 10 | Clustering algorithms | 7 | Knowledge on K-mean, |
| 11 | Difference between L1 & L2 regularizations | 6 | Explained |
| 12 | Boost Strapping Vs Bagging Algorithms | 7 | Answered |
| 13 | Difference between Normalization and Standardization | 8 | Explained |
| 14 | How to identify Hyper Parameters | 7.5 | Explained the grid search technique |
| 15 | Hands-on experience with Word Embedding Techniques like TF-IDF, word2vec,  Glove, BERT, and embedding models like ada. | 4 | Limited exposure on NLP concept |
| 16 | Basic understanding of LLMs / GPT-based models and their applications. | 0 | Not worked on |
| 17 | What is K-fold validation | 8 | Explained |
| 18 | Sampling techniques | 6 | Partially Explained |
| 19 | Experience in Databricks / PySpark for scalable data processing | 5 | Mostly worked on python / databricks (on AWS) |
| 20 | What are the data pre-processing steps? What are different data patterns? | 6 | 1. Data spread 2. Trend, Seasonal, Irregular, |
| 21 | Coding skills – python and SQL | 8 | Python – 8/10, SQL – 9/10 |