<u>Instagram</u> | <u>GitHub</u> | <u>Medium</u> | <u>YouTube</u>

If you are seriously interested in AI/ML

- 1. **Programming**: Python (Core & Data Science), SQL, and deep learning frameworks like PyTorch/TensorFlow
- 2. **Statistics & Probability**: Hypothesis testing, Mean, Median, Mode, Variance, Standard Deviation, Quartiles, Bayesian vs Frequentist, Correlation vs Causation p-values, Confidence intervals, Type I/II errors, Random sampling, Stratified sampling, Sampling bias, Probability distributions (normal, binomial, Poisson), Conditional probability, Bayes' theorem
- 3. Data Wrangling & EDA: Cleaning, transforming, and preparing data for analysis
- 4. Data Visualization: Tools like Matplotlib, Seaborn, Tableau or Power BI
- 5. Feature Engineering: Creating and selecting meaningful features from raw data
- 6. Machine Learning
 - Supervised Learning: Regression (Linear/polynomial regression, regularization (Ridge, Lasso), logistic regression), Classification (Decision Trees, Random Forests, SVM, Naive Bayes, K-NN), Ensemble Methods(Bagging, Boosting (XGBoost, LightGBM), Stacking)
 - Unsupervised Learning: Clustering (K-means, Hierarchical Clustering, DBSCAN),
 Dimensionality Reduction (PCA, t-SNE etc)
 - **Time Series Analysis:** Forecasting Methods (ARIMA, seasonal decomposition, exponential smoothing), Time Series Features (Trend, seasonality, stationarity, autocorrelation), Modern Approaches (Prophet, LSTM for time series)
- 7. **Deep Learning Fundamentals:** Neural Networks (Perceptrons, backpropagation, activation functions), Architectures (CNNs (computer vision), RNNs/LSTMs (sequences), Transformers), Training Techniques (Gradient descent variants, batch normalization, dropout)
- 8. Model Evaluation & Selection:
 - Metrics: Accuracy, precision, recall, F1-score, ROC-AUC, RMSE, MAE
 - Validation: Train/validation/test splits, cross-validation, stratified sampling
 - Model Selection: Bias-variance tradeoff, overfitting/underfitting, hyperparameter tuning
 - Performance Analysis: Learning curves, confusion matrices, feature importance
- 9. Big Data Tech: Spark, or cloud data platforms (AWS, Azure, GCP)
- 10. Database Knowledge: Relational (SQL) and NoSQL databases

- 11. **Mathematics**: Linear algebra, calculus (basics relevant to ML/LLM algorithms), Optimization (Gradient descent, local/global minima, cost functions), Understanding floating-point arithmetic, numerical stability
- 12. **Communication**: Ability to explain findings clearly to technical and non-technical audiences
- 13. Domain Knowledge: Understanding of the specific industry you work in
- 14. Version Control: Git and collaboration workflows
- 15. Cloud Computing: Deploying models and workflows on cloud infrastructure
- 16. **Experimentation**: A/B testing and designing experiments
- 17. MLOPs: Model deployment, monitoring & lifecycle management
- 18. **GenAI**: NLP & Transformer fundamentals, Working with LLMs, Prompt engineering, RAG, Vector Databases, Embedding Models, Multimodal, Model Optimisation Agentic AI, Finetuning open-source models, evaluation