Roadmap to be a ML Engineer

1. Basic Knowledge of Computer
2. Basic Programming Language Python / C / C++ / Java
3. Data Structure and Algorithm
4. SQL
5. Maths
    - Linear Algebra
    - Calculus
    - Probability
    - Statistics
6. *Machine Learning Fundamental*
7. Specific Topic NLP / CV
8. Apply Basics of ML in Your Specialization NLP
9. Deep Learning Fundamental
10. Be Expert on Specific topic (NLP, CV) with Deep Learning
11. ML OPs for Deploy

Machine learning Fundamental

1. Environment Setup
2. Numpy (prerequisite)
3. Pandas (prerequisite)
4. Matplotlib (prerequisite)
5. Types of Machine learning
6. Classification vs Regression
7. Train , Test and Evaluation
8. Again Introduction with Supervised algorithm / overview of Algo

Algorithm Part - Supervised

9. Linear regression
10. Performance
11. Overfit & Underfit
12. Bias variance
13. Bias variance trade off
14. Outliers
15. Residual - cost & loss function
16. Evaluation for Regression
    a. MAE
    b. MSE
    c. RMSE
    d. R^2 Score

17. Gradient Descent
18. Types of Gradient Descent
19. Batch gradient descent
20. Scotestic gradient descent
21. Mini batch gradient descent
22. Linear regression with Gd
23. Multiple Linear Regression
24. Multiple linear regression with GD
25. Polynomial regression
26. Ridge regression
27. Lasso Regression
28. Elastic Net Regression
29. SGD Regression
30. Logistic Regression
31. Evaluation For Classification
    a. Confusion Matrix
    b. Accuracy score
    c. Precision
    d. Recall

    e. F1-Score
    f. ROC curve and AUC
    g. Classification Report
32. KNN
33. Conditional Probability - naive bayes
34. Bayes Theorem - naive bayes
35. Naive Bayes
36. SVM
37. SVR
38. SVC
39. Decision Tree for classification
40. Decision Tree for regression
41. Ensemble
42. Bagging
43. Boosting
44. Stacking
45. Blending
46. Random Forest classification
47. Random Forest Regression
48. Adaboost
49. Gradient Boosting
50. XGboosting
51. LightBoosting
52. Catboosting
53. Introduction with UnSupervised algorithm / overview of Algo

Algorithm Part - UnSupervised

54. K-means clustering
55. Evaluation for unsupervised
    a. Silhouette score
    b. Davies-bouldin index

56. Hierarchical clustering
57. DBSCAN
58. Gaussian Mixture Models

Dataset Handling and Others

59. Workflow of a project.
60. Time complexity of ML algorithm

Data Cleaning and Preprocessing

61. Handling the missing value
62. Mean, median, variance, standard deviation
63. 5 summary of statistics - outliers
64. Detect Outliers
65. Handle the Outliers
66. Handle the imbalance Data
67. What is Data distribution , Normalization

68. Handle the skewed data
69. Detect and Handle the anomaly
70. Anomaly vs Outliers
71. Handling Date and Time

Feature Engineering

72. Multicollinearity concept
73. Feature selection - matrix correlation
74. Removing Multicollinearity
75. Feature Importance by SHAP, LIME
76. Feature Extraction by PCA
77. Encoding
78. Binning and Discretization
79. Scaling
80. Column Transform
81. Log Transform
82. Pipe Line

Hyperparameter Tuning

83. GridsearchCv and RandomizedSearchCv

End To End Project

84. Mental Health - Kaggle Competition
85. House Price - Kaggle Competition
86. Loan Approved - Kaggle Competition

End to End Project for CV

87. Movie Recommendation
88. Sentimental Analysis