I already explained ===== how to connect chatgpt API

Take some cloud account

Azure account ======

Azure for students ====== no credit card details ==== 100$ free

College email id ====

Azure for non students ==== credit /debit ===== 200$ free ==== deactivate account

How to create azure free account ===== email credit/debit

Azure ML services ==== keys python

By 26th ========

Linear regression ========= OLS

Logistic regression ======= sigmoid

Decision Tree ========== Entropy /gini

Naïve bayes ========= Probability

KNN ========== distance metric

**Classification Metrics**

1. **Confusion matrix**
2. **Accuracy**
3. **Error rate**
4. **Sensitivity**
5. **Specificity**
6. **Precision**
7. **Recall**
8. **F1-score**
9. **Kohens-kappa**
10. **ROC-AUC curve**

**Age gender income cr y**

**30 M 50k fair yes**

**You develop the model ====== train data to the model**

**{30,M,50K,fair} ============ yes No**

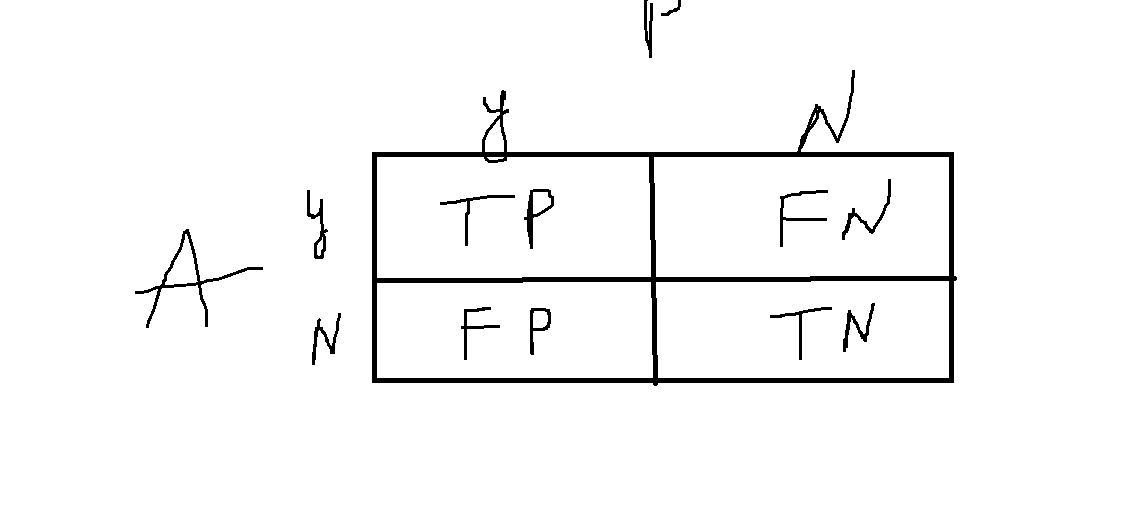
**How many total combinations**

**Actually yes ,model predicted as No**

**Actually yes, model predicted as yes**

**Actually No , model predicted as No**

**Actually No, model predicted as yes**

****

**In a data originally 50 yes are there , 50 No are there**

**Model is predicted 40 yes and 60 No**

**Create the Confusion matrix**

**Tp=0 tn=99 99+0 = 99**

**Tomorrow you have a class 9am**

**Azure ml studio Mlops**

**Confusion matrix**

**TP TN FP FN**

**FP ===== > model is predicting yes but it is wrong**

**Actually no**

**Accuracy**

**Error rate**

**Class imbalance problem:**

**Yes No**

**100**

**90 10**

**80 20**

**95 5**

**Classes are imbalanced accuracy will fail**

**How to deal with class imbalance problem**

**99 yes 1 No**

**Cancer patient detection**

**Health related rare**

**99 yes 1 No ======== to make the balance**

**Add data**

**1 -=========== 99**

**99 ============ 1**

1. **Up sampling**

**You will concentrate on lower label observation**

**You will increase those observations**

**No cases**

**30 M 50k No**

**31 M 50k No**

**30 F 50k No**

1. **Down sampling**

**99:1**

**2000 8000**

**4000 4000**

**Red green yellow**

**200 200 200 600**

**LLM fine tuning**

**Google**

**Spam reviews data**

1. **Website**
2. **Telephone**
3. **Google maps**

**285 reviews 12Yes NO**

**50 reviews 13 277 NO run the model**

**5 1 yes 4NO**

**TPR= True positive rate = Tp/p ========= sensitivity**

**TNR= True negative rate= Tn/n ======== specificity**

**Sensitivity =0 specificity= 99 accuracy =99**

Precision and Recall

| **X** | **Y** | **Y\_pre** |
| --- | --- | --- |
|  | **Yes** | **No** |
|  | **Yes** | **Yes** |
|  | **Yes** | **Yes** |
|  | **No** | **Yes** |
|  | **No** | **No** |

**Qn) How many actual positives are there = 3**

**Qn ) How many model predicted positives are there =3**

**Qn) How many model positive predictions are matching with actual positives=2**

| **Classes** | **Yes** | **No** | **Total** | **Recognition** |
| --- | --- | --- | --- | --- |
| **Yes**  **No** | **90(TP)**  **140(FP)** | **210(FN)**  **9560(TN)** | **300**  **9700** | **30.00%**  **98.56%** |
| **Total** | **230** | **9770** | **10,000** | **96.40** |