

MLFlow

Start at 9:05

Experiment Tracking & Data Management

→ Git → Store versions of code

ML model →

- data
- data processing
 - outlier removal
 - Null values
- Split into test & train
 - encoding
- Choose a model
- Hyper tune parameter
- evaluation metric / loss function

→ loan_app_dt_model1.py
- model2.py } → File management

→ Team used to maintain tables/dataset of all iteration you have performed

model version	Param	eval	Prec	recall
V ₁				
V ₂				
V ₃				
V ₄				

→ • accuracy →
 → 98% p 80% acc
 → 90% p 85% acc

Precision → 95%
 86%
expectation mismatch

→ Miyatra →

Fashion → Mens T-shirt
 → ()
 89%

Sports wear
 → ()
 → 69%

→ Share all attempts I have
 with diff approach

→ So 2nd party can
 reproduce

⇒ • Organise →
 • Reproduce →
 • Logging the results

→

1	1
---	---

 → LR → dataset 80/20 hyperparam
 →

acc	1%	recall
-----	----	--------

⇒ M2Flow

→ • Store data → if it changes frequently
 • BDA →

1

⇒ RF Classifier

→ max-depth = [2, 4, 6, 8, 9,] ←

n-est = [10, 20, 30, 40] ←

2 x 10

2 x 20

2 x 30

or either
you run
all
model
combination →

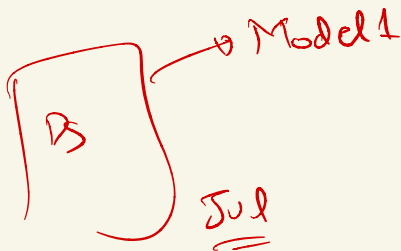
→ RF

KNN

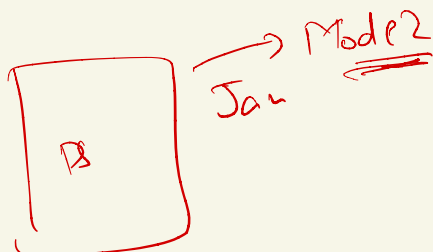
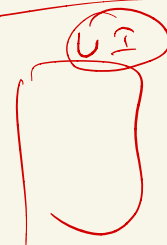
DT

→ Best RF
Hyper para
L

→



→



→



