E-Prescription Recommendation

Internship at Prixa.ai
Dave Danadiva Agusta Peerera (12006002)
dave.peerera@student.sgu.ac.id

INTRODUCING

01

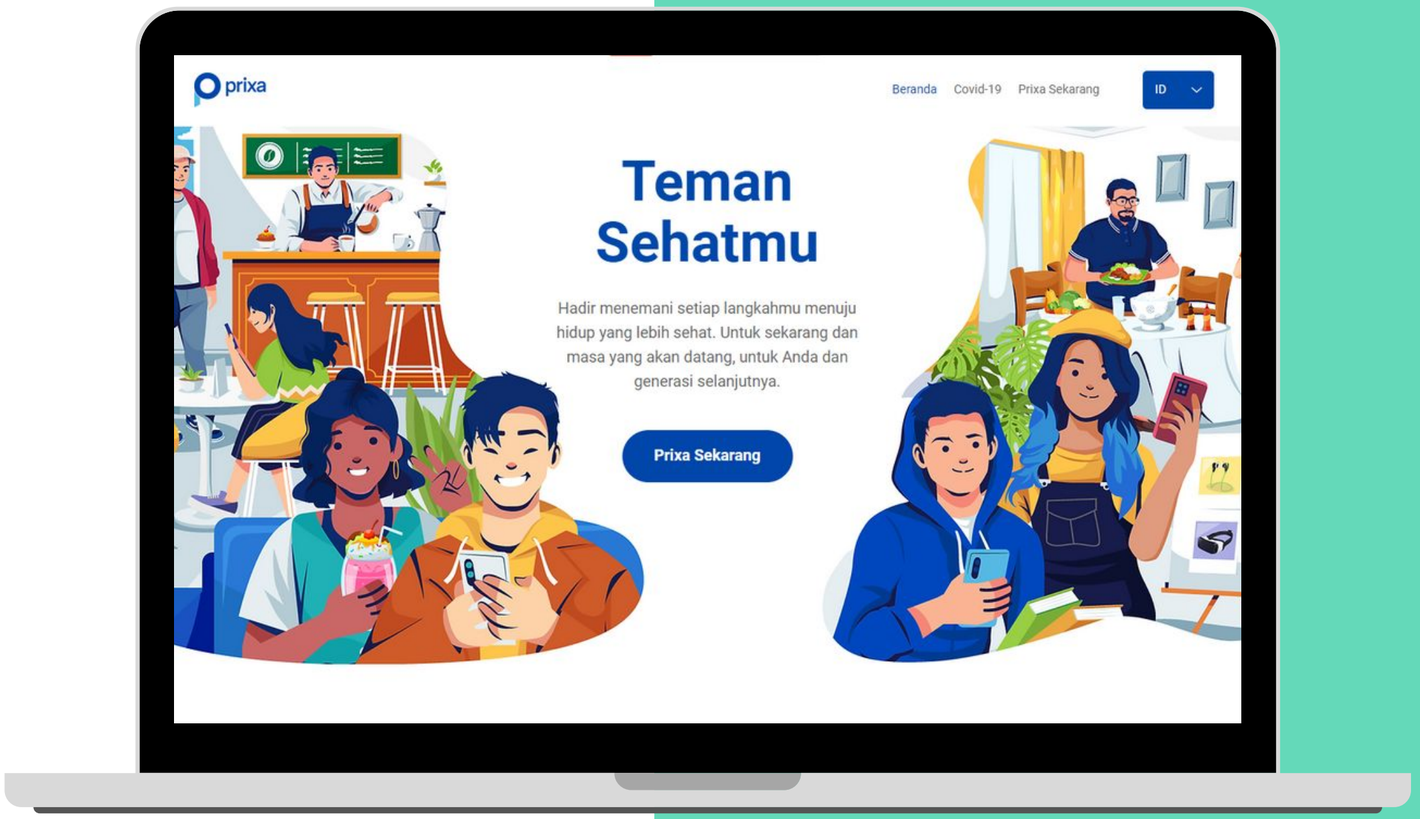
VISION

To enable a promising future for you and the next generation

02

MISSION

Humanize healthcare using data and technology

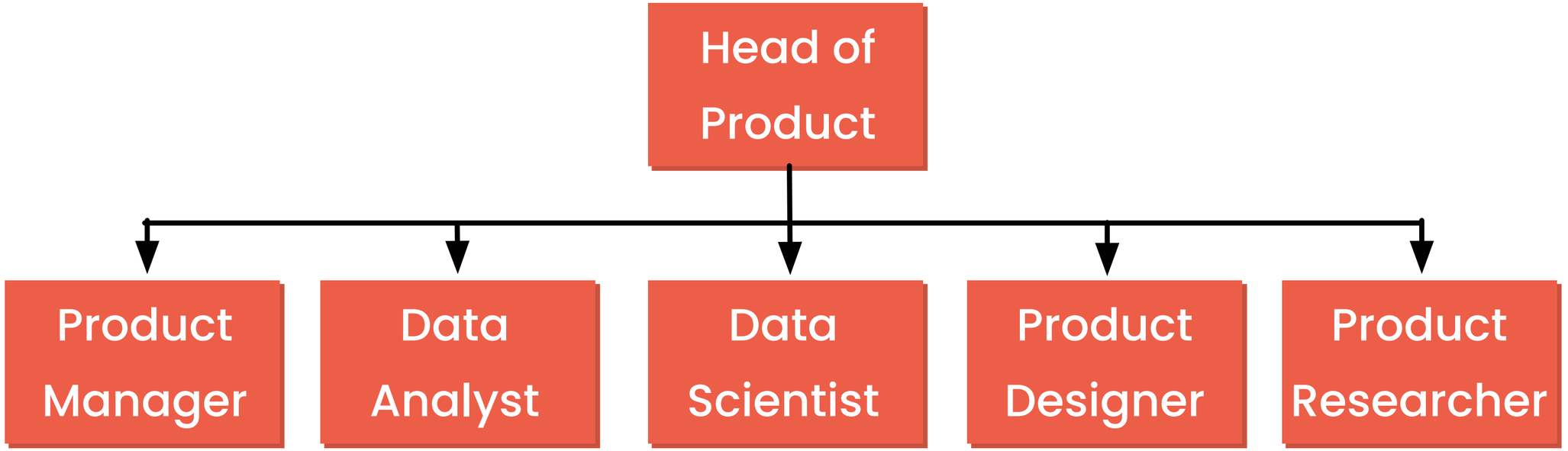


WEB-APP SERVICES

- Symptom Checker 01
- Teleconsultation 02
- Lab Test 03
- Pharmacy Delivery 04
- Doctor Appointment 05

For more info visit: prixa.ai

PRODUCT TEAM STRUCTURE



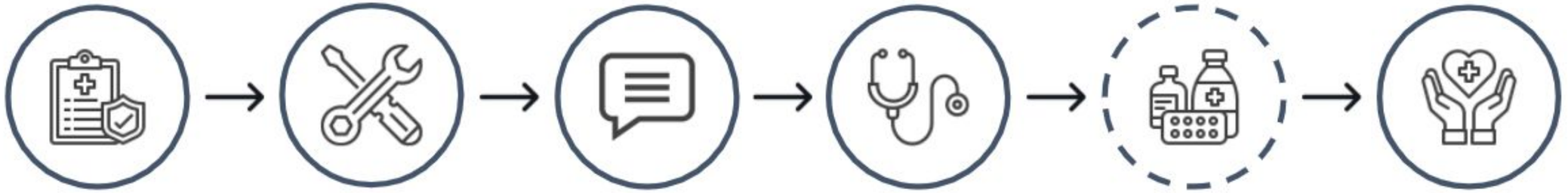
E-Prescription Recommendation

Goals

Develop and evaluate a medicine drug recommendation system that can assist doctors in making decision faster, more precise, and accurate prescriptions for patients with certain diseases.

BACKGROUND

Doctor spent approximately 23 minutes for each teleconsultation session



±1 minute
Prepare for consultation

±2-3 minutes
Uncover patient's current symptoms

±10 minutes
Interview patient

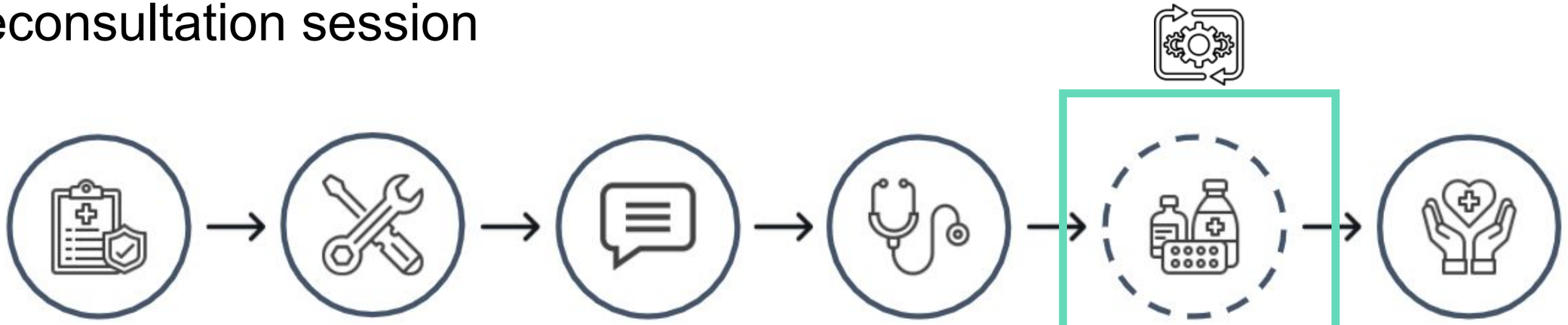
±2-5 minutes
Educate & give recommendations

±6-10 minutes
Issue Prescription

±2-5 minute
Resolve teleconsultation

BACKGROUND

Doctor spent approximately 23 minutes for each teleconsultation session



±1 minute
Prepare for consultation

±2-3 minutes
Uncover patient's current symptoms

±10 minutes
Interview patient

±2-5 minutes
Educate & give recommendations

±6-10 minutes
Issue Prescription

±2-5 minute
Resolve teleconsultation

TOOLS

Python 3.7 & Python 3.10

Google Colaboratory IDE

Database

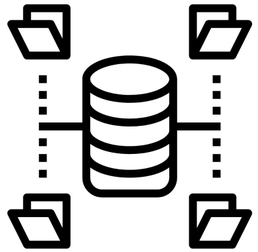
Visual Studio Code IDE

Visualization

METHODOLOGY

PROJECT STAGES

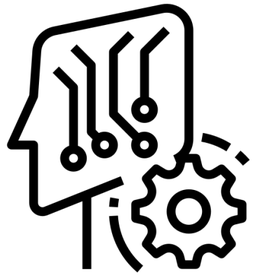
01 DATA GATHERING & CLEANING



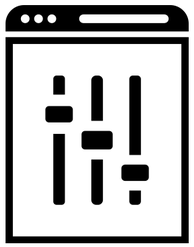
02 EXPLORATORY DATA ANALYSIS

03 DATA PREPROCESSING & SPLITTING

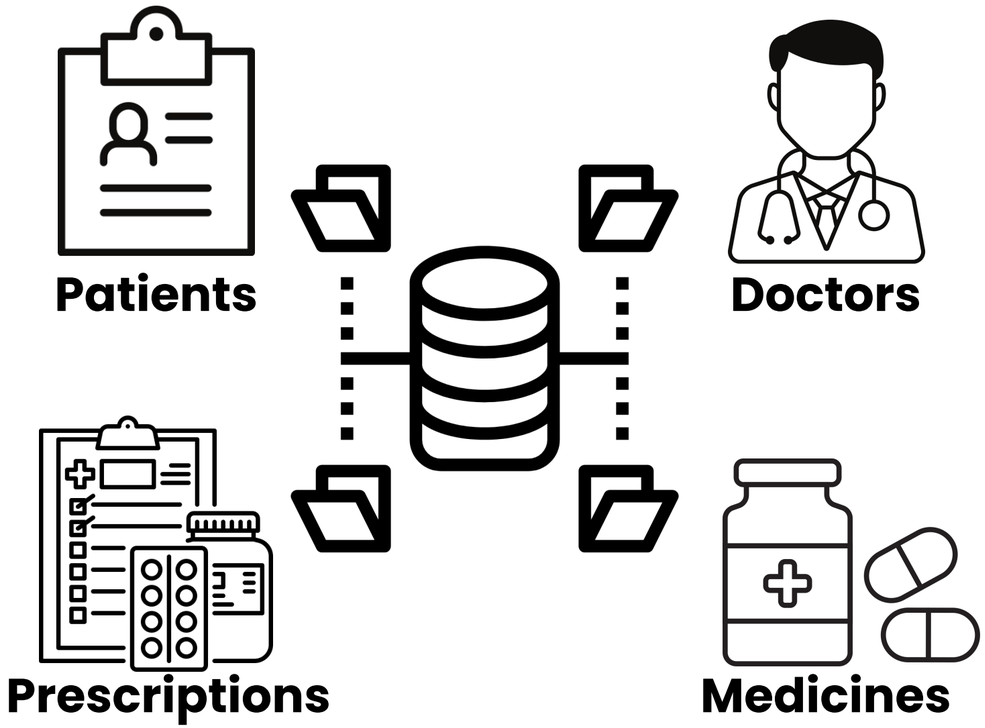
04 AI MODEL DEVELOPMENT & EVALUATION



05 MODEL REFINEMENT & SELECTION



06 MODEL DEPLOYMENT



DATA GATHERING & CLEANING

- Gather data 01
- Convert data file format 02
- Filter out dummy data 03
- Filter out rubbish data 04

STAGE 2

METHODOLOGY

EXPLORATORY DATA ANALYSIS

01

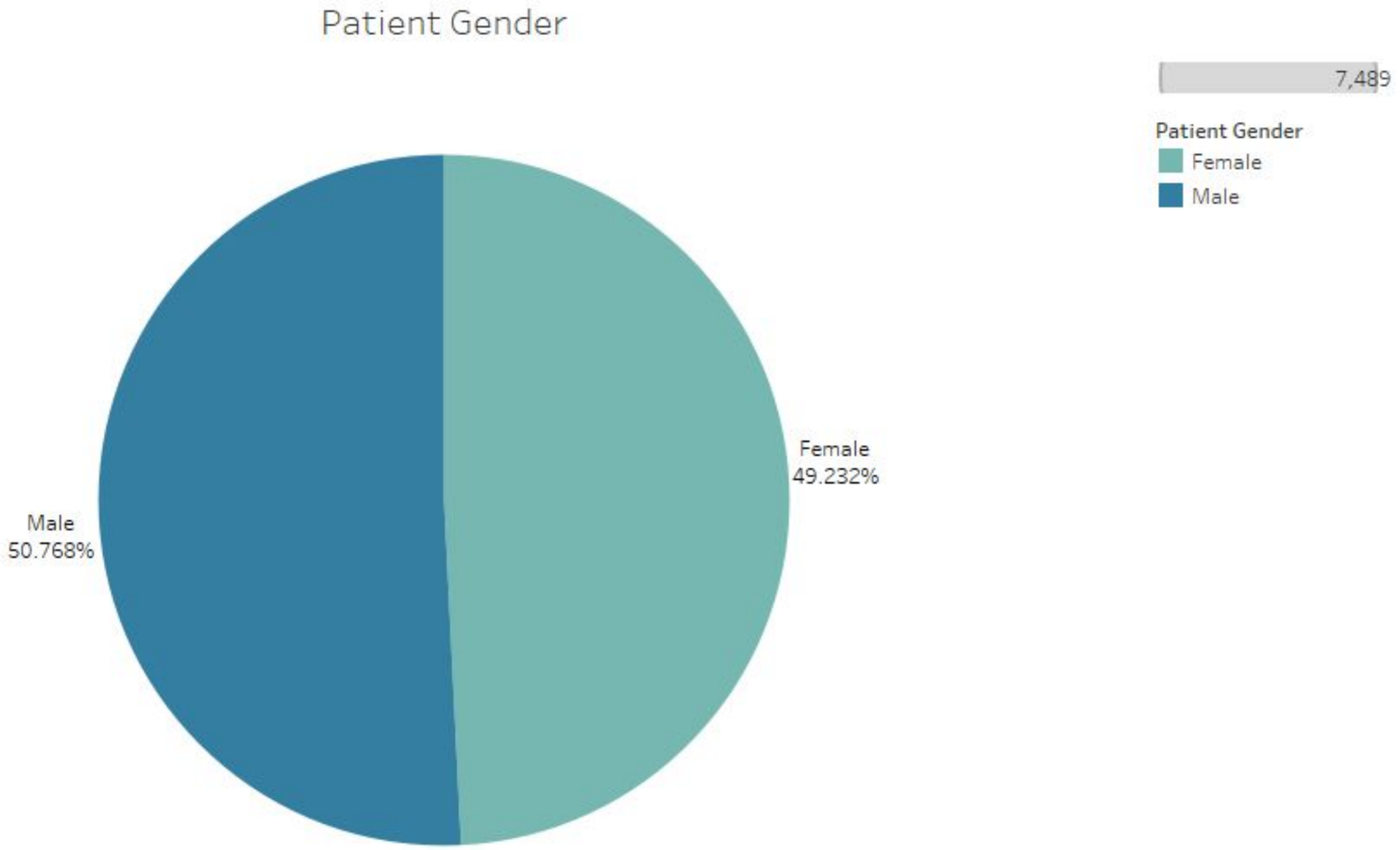
Explore Data gathered to gain insight and obtain correlation between features in the dataset

02

Exploratory data analysis will be divided into patients and prescriptions part and medicine part

METHODOLOGY

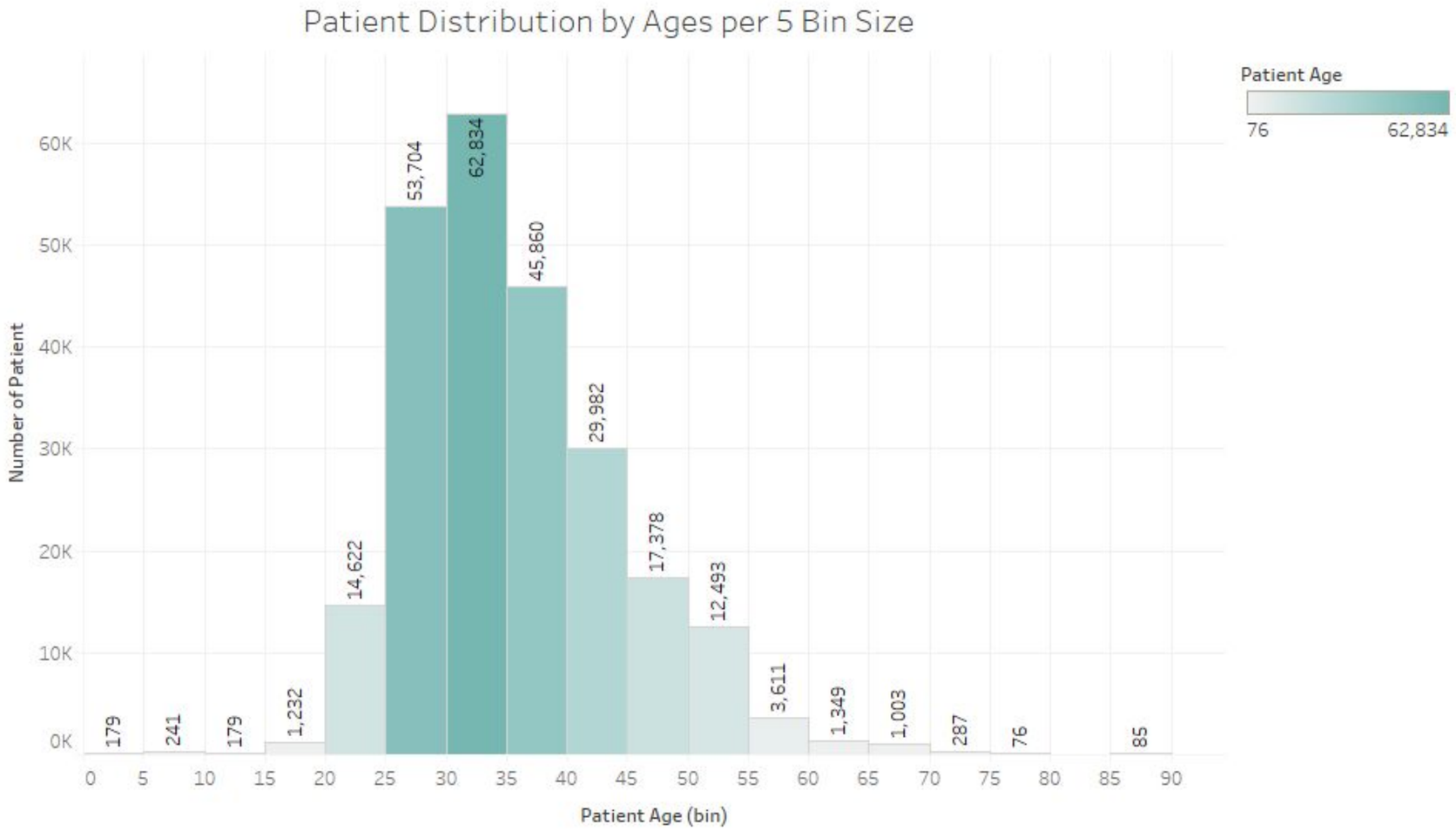
EXPLORATORY DATA ANALYSIS



PART 1: PATIENTS & PRESCRIPTIONS

METHODOLOGY

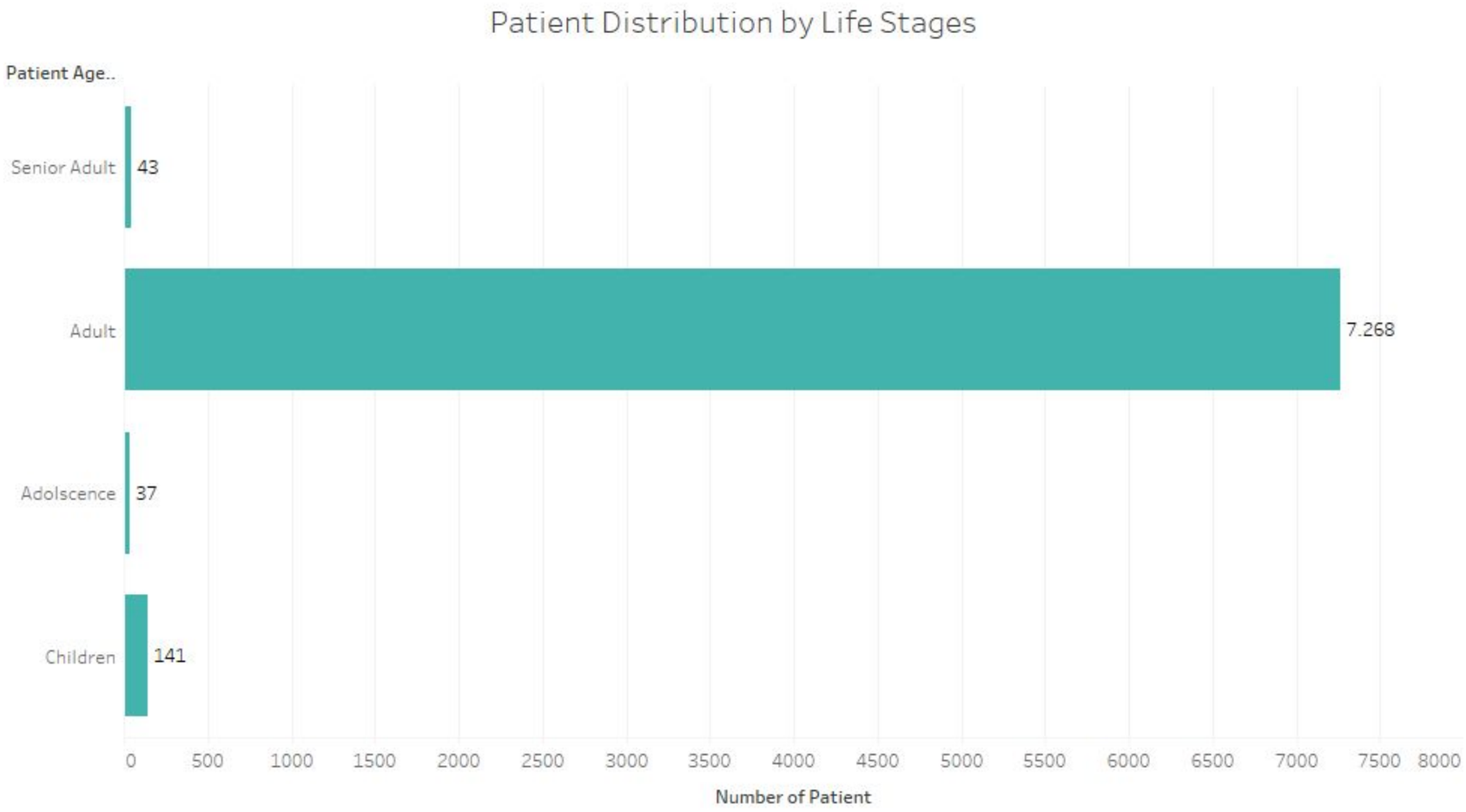
EXPLORATORY DATA ANALYSIS



PART 1: PATIENTS & PRESCRIPTIONS

METHODOLOGY

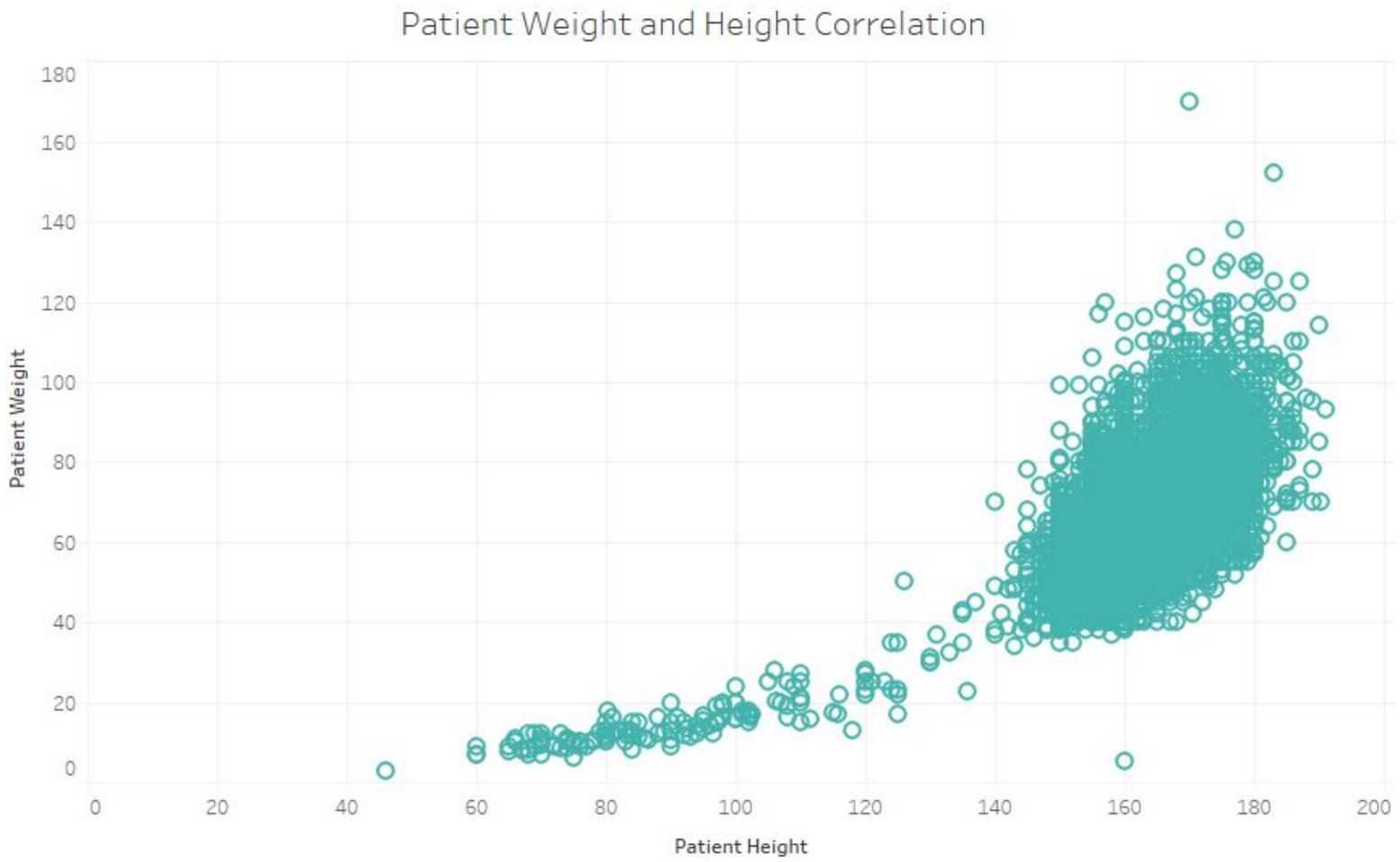
EXPLORATORY DATA ANALYSIS



PART 1: PATIENTS & PRESCRIPTIONS

METHODOLOGY

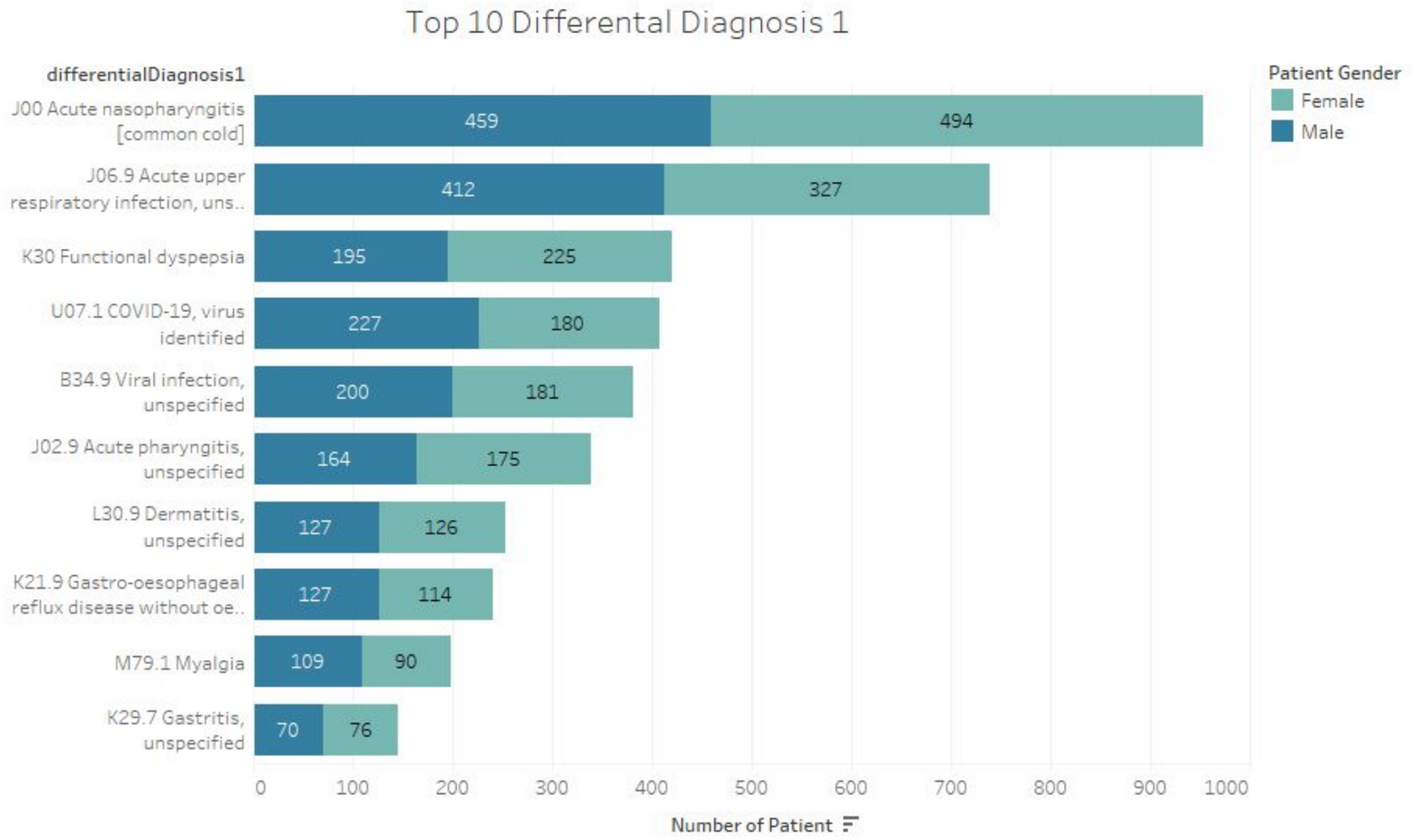
EXPLORATORY DATA ANALYSIS



PART 1: PATIENTS & PRESCRIPTIONS

METHODOLOGY

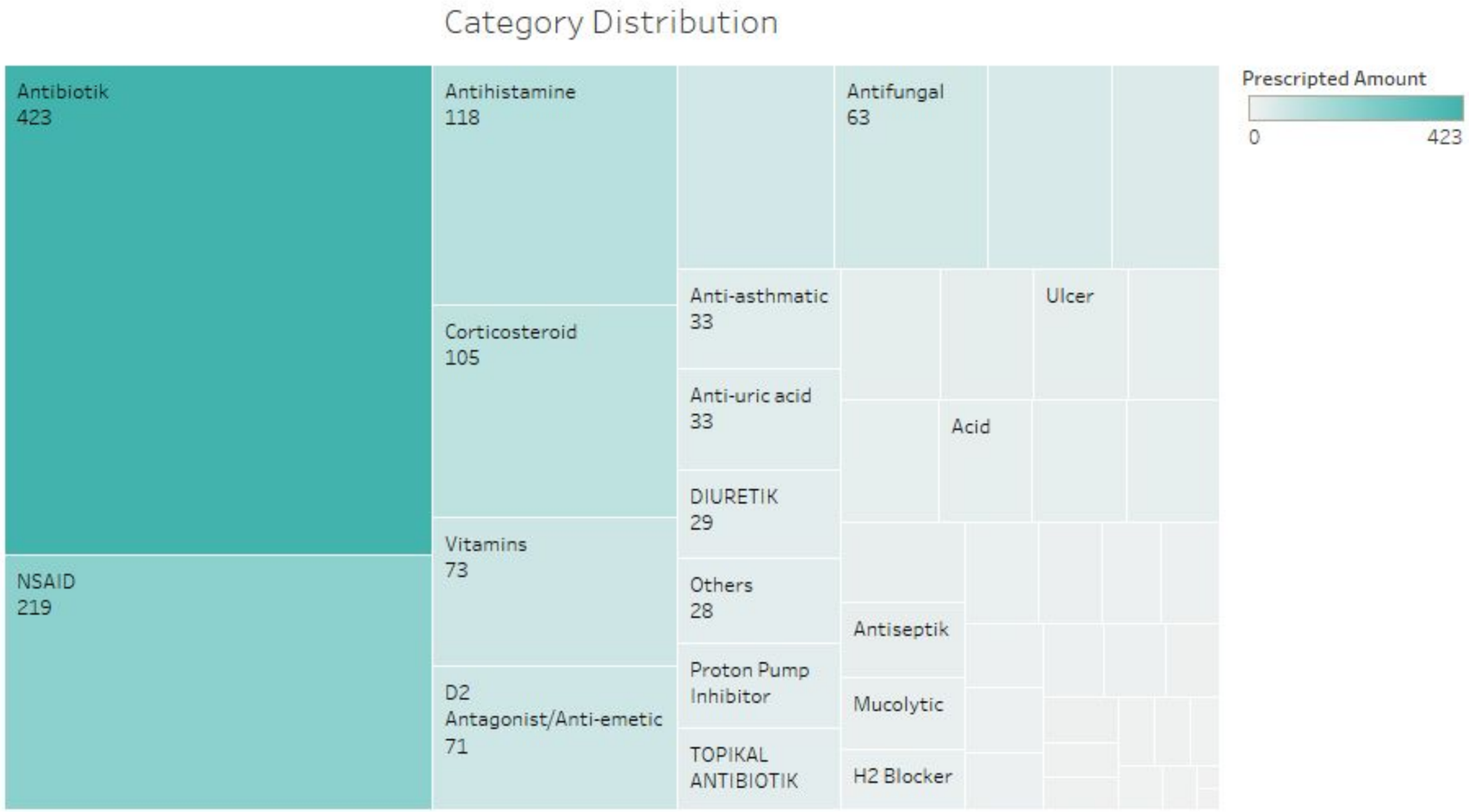
EXPLORATORY DATA ANALYSIS



PART 1: PATIENTS & PRESCRIPTIONS

METHODOLOGY

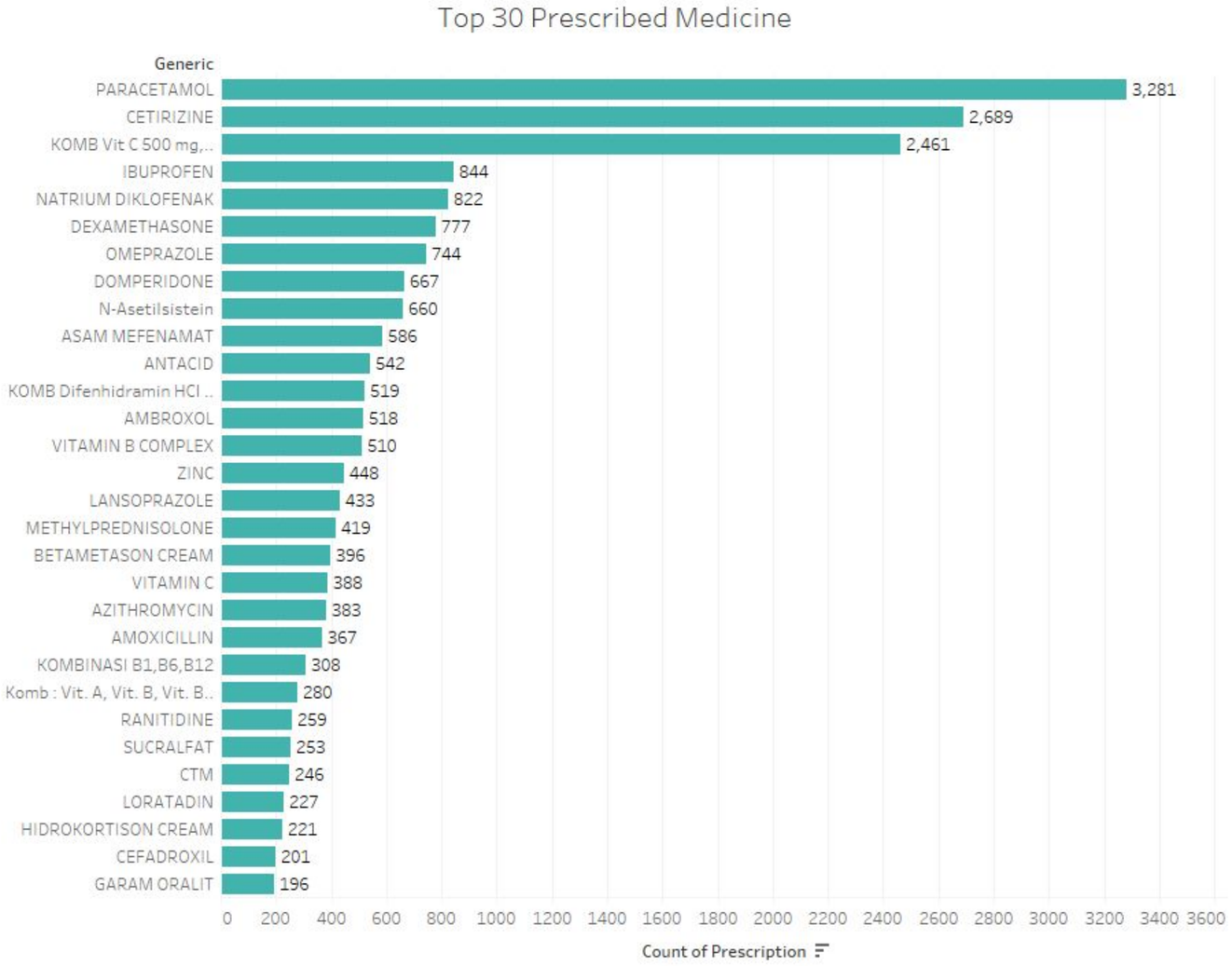
EXPLORATORY DATA ANALYSIS



PART 2: MEDICINES

METHODOLOGY

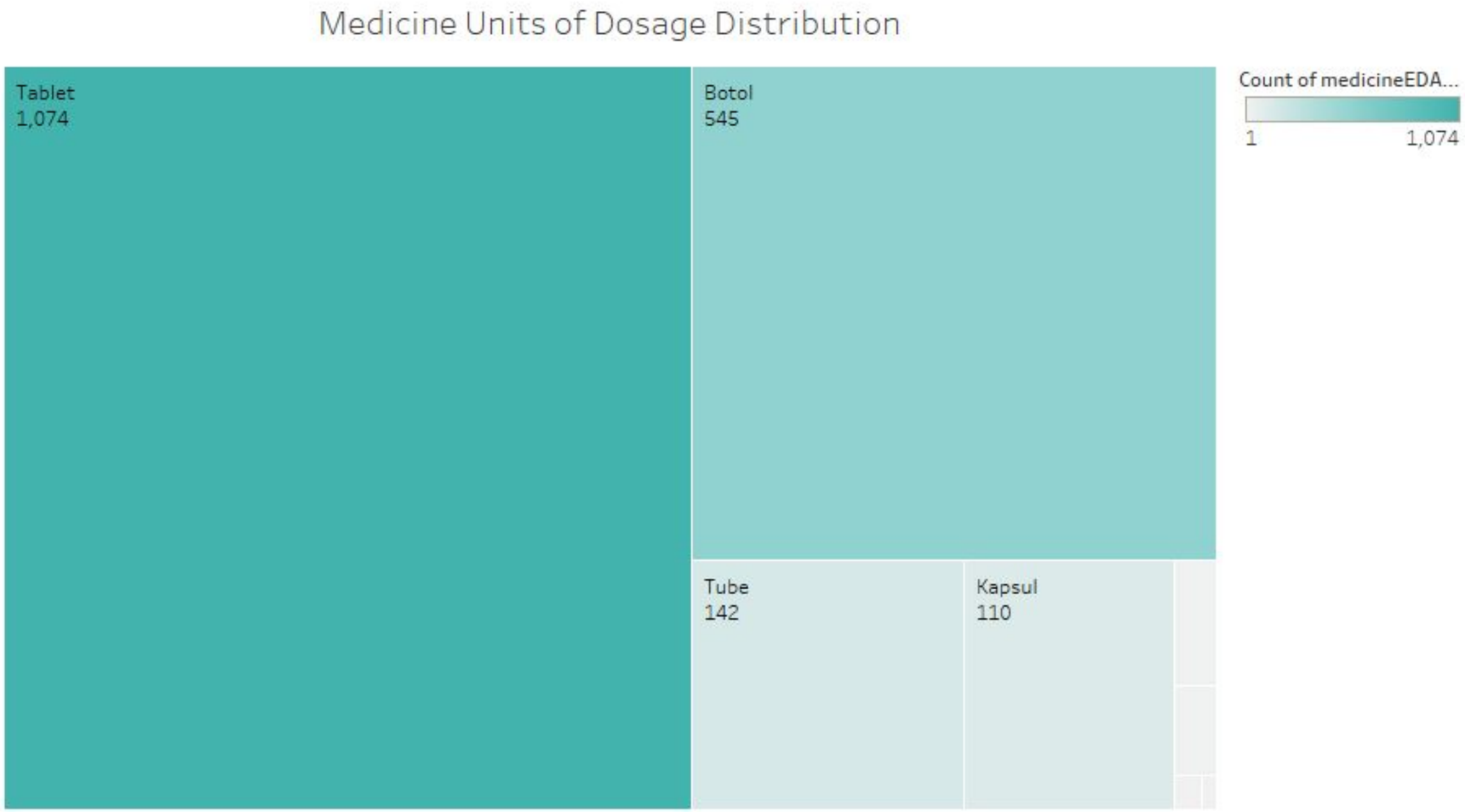
EXPLORATORY DATA ANALYSIS



PART 2: MEDICINES

METHODOLOGY

EXPLORATORY DATA ANALYSIS



PART 2: MEDICINES

METHODOLOGY
DATA PREPROCESSING & SPLITTING

STAGE 3

Prepare and ensure the quality of the data before used for model development

Missing Value

Resolve by skewness of data

Numeric Data

- if skewness ≈ 0; use mean
- if skewness < 0 or > 0; use median

Object Data

- use mode

Default Medicine

Map out top 50 ICD-10 with their popular prescribed medicine into json file.

- ICD name as key
- medicine & medicine count as value

METHODOLOGY

DATA PREPROCESSING & SPLITTING

Binning Patient Age

Age Range	Life Stage
0 -12	Children
13 - 18	Adolescence
19 - 59	Adult
≥ 60	Senior Adult

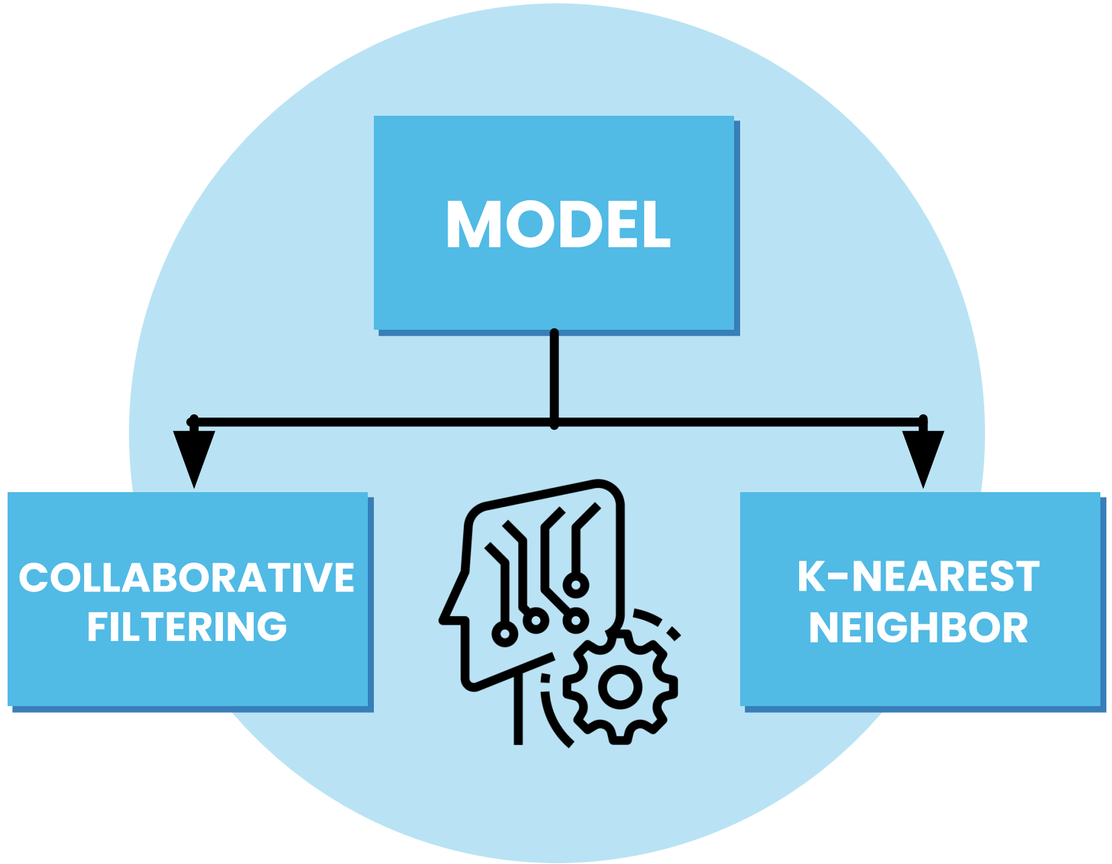
Binning Patient BMI

Age Range	BMI Category
< 18.5	Underweight
18.5 - 24.9	Normal
≥ 25	Overweight

METHODOLOGY

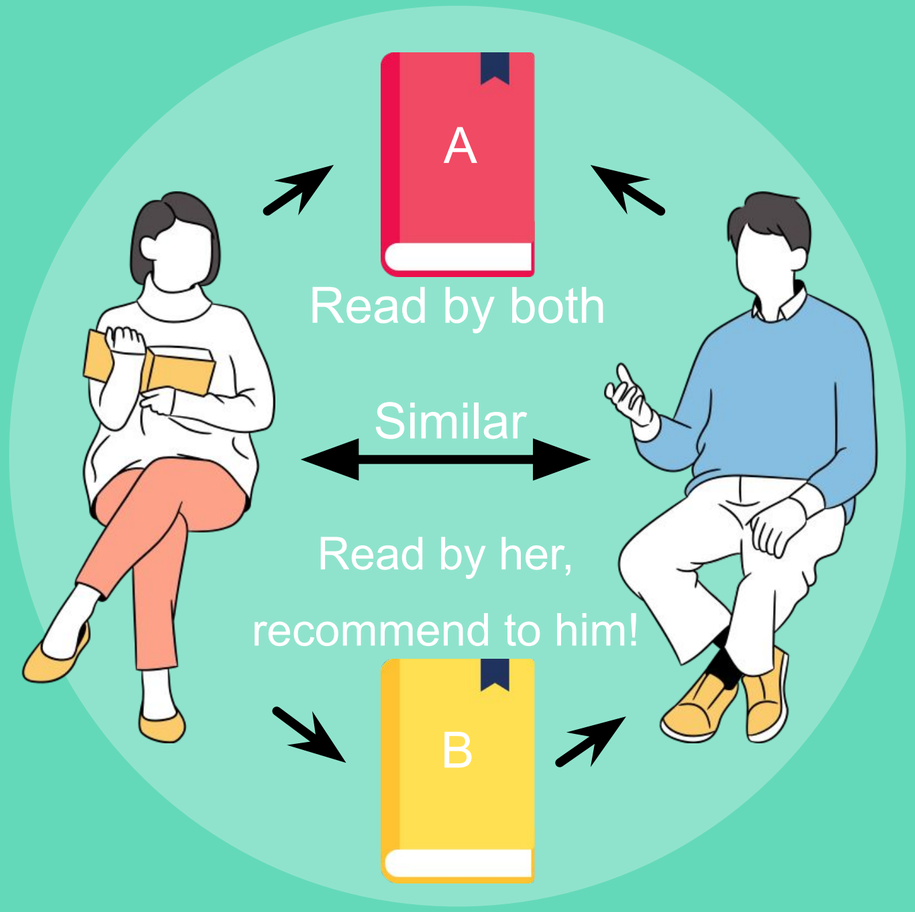
STAGE 4

AI MODEL DEVELOPMENT & EVALUATION



MODEL 1: COLLABORATIVE FILTERING

"If a person A has the same opinion as a person B on an issue, A is more likely to have B's opinion on a different issue than that of a randomly chosen person."

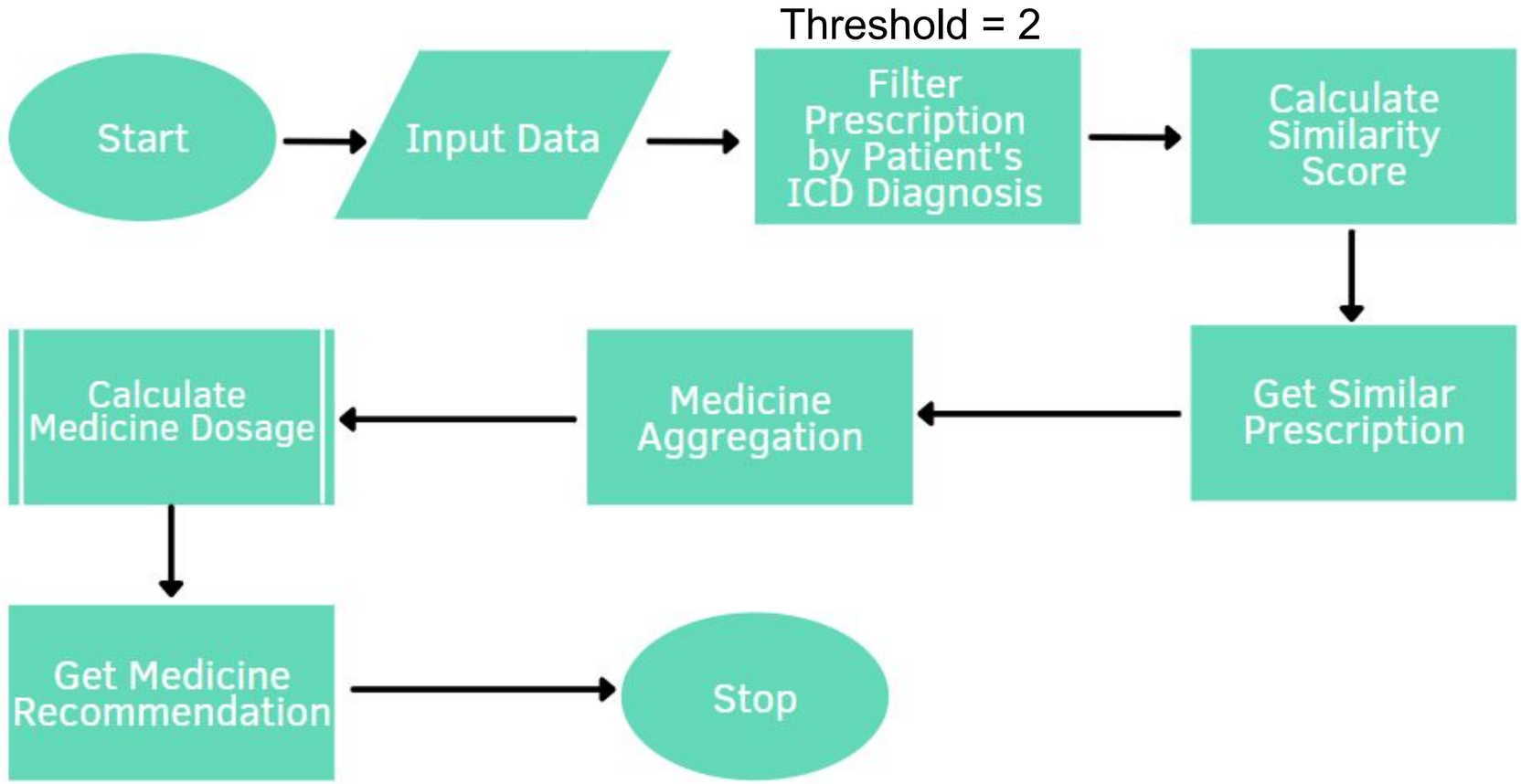


MODEL 1: COLLABORATIVE FILTERING

SCORING METHOD

Parameter	Range	Description	Score
ICD-10 Diagnosis	-	If all of the ICD-10 diagnoses match with the other patient, it will give a full score. If it only matches partially, ex. 2 out 5; hence it will be ⅖.	4
Age Category	Difference: 1 - 2 [-0.25 Score] Difference: 2 - 4 [-0.50 Score] and so on [+ -0.25 Score]	For a different of more than one age category to the active patient, 0.25 will be subtracted from 1 (=100% similarity) for each additional age bin	1
BMI Category	-	If the BMI category matched, score is 1, else 0.	4
Gender	-	If gender matched score is 1, else 0.	1
Total Score			10

MODEL 1: COLLABORATIVE FILTERING



ALGORITHM FLOW

MODEL 1: COLLABORATIVE FILTERING

INPUT DATA

Prescription	Prescription Pivot	Patients	Medicines	Medicine Prescriptions
Prescription ID	Patient ID	Patient ID	Medicine Brand	Frequency
Differential Diagnosis	Medicine Brand	Patient's Age Category	-	FrequencyDd
-	-	Patient's BMI Category	-	Timing
-	-	Gender	-	Duration
-	-	-	-	Amount

MODEL 1: COLLABORATIVE FILTERING

OUTPUT DATA: MEDICINE LIST


```
{
    "SURBEX T (KOMB Vit C 500 mg, Niasinamida 100 mg, Kalsium Pantotenat
20 mg,Vit. B1 15 mg, Vit. B2 10 mg,Vit. B6 5 mg, Vit. B12 4 mg) TAB": 124,
    "CETIRIZINE 10 MG Tablet": 78,
    "PARACETAMOL 500 MG Tablet": 77,
    "AMBROXOL 30 MG Tablet": 48,
    "Acetylcystein 200mg Tablet": 33,
    "SANADRYL DMP (Dekstrometorfan HBr 10 mg, Difenhidramin HCl 12,5 mg,
Ammonium Cl 100 mg, Na Sitrat 50 mg, Mentol 1 mg/ 5ML) SYRUP, BOTOL
120ML": 23,
    "DEXAMETHASONE 0.5 MG Tablet": 20,
    "PARACETAMOL 600 MG Tablet": 19,
    "CETIRIZINE 5 MG/5 ML Botol": 18,
    "ELKANA Syrup Botol 60ML": 18
}
```

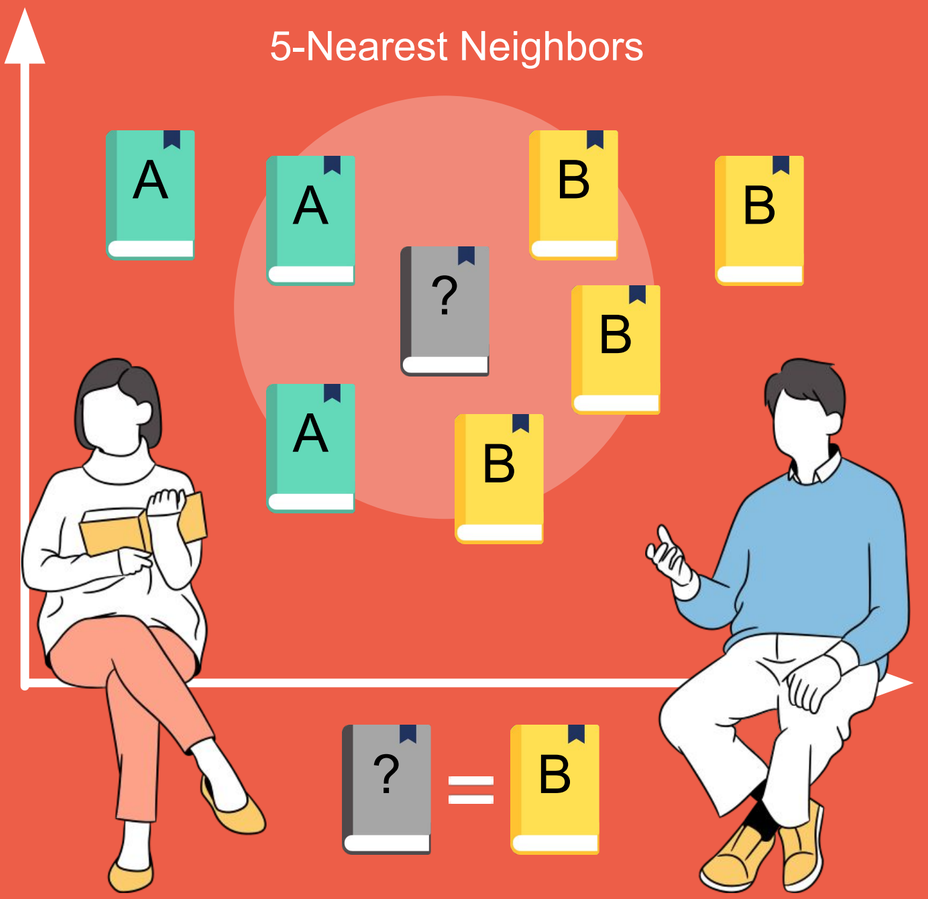
MODEL 1: COLLABORATIVE FILTERING

OUTPUT DATA: MEDICINE DOSAGE

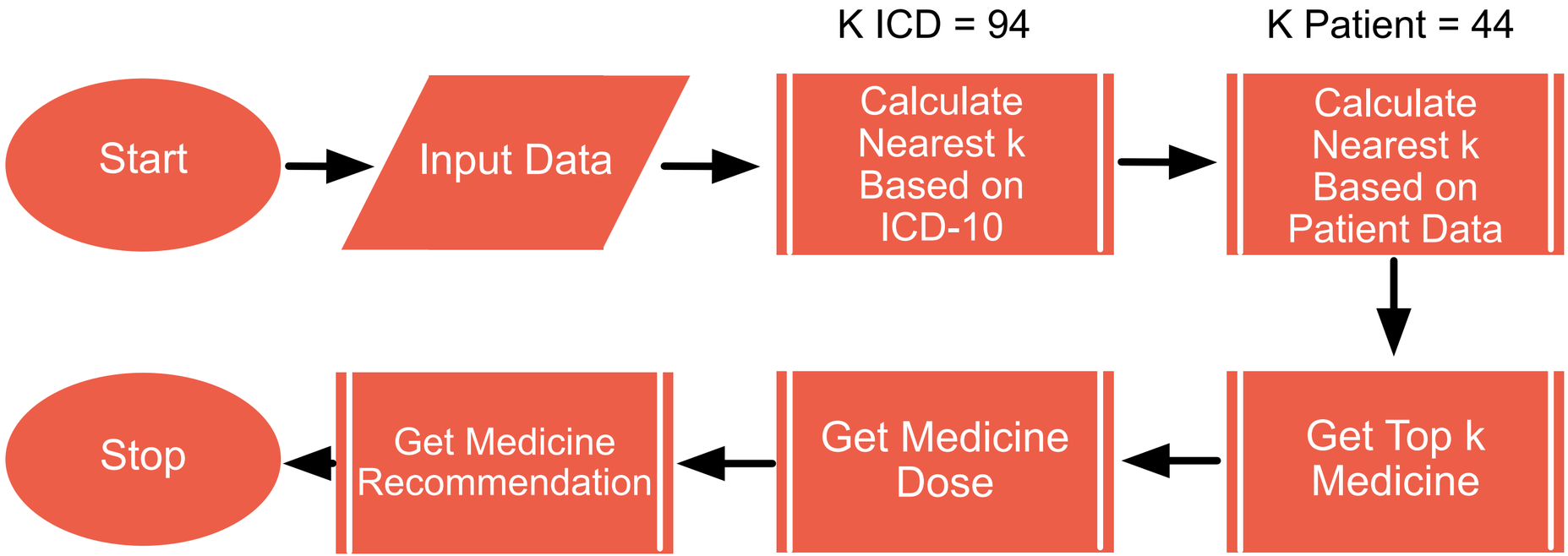

```
"SURBEX T (KOMB Vit C 500 mg, Niasinamida 100 mg, Kalsium Pantotenat 20
 mg,Vit. B1 15 mg, Vit. B2 10 mg,Vit. B6 5 mg, Vit. B12 4 mg) TAB": {
    "frequency": 1.0,
    "frequencyDd": 1.0,
    "timing": "Setelah Makan",
    "duration": 5.0,
    "amount": 5.0
}
```

MODEL 2: K-NEAREST NEIGHBOR

"Unknown Data point that is in a certain neighborhood will likely be similar with the most data point in that neighborhood"



MODEL 2:
K-NEAREST NEIGHBOR



ALGORITHM FLOW

MODEL 2:
K-NEAREST NEIGHBOR

INPUT DATA

Prescription	Prescription Pivot	Patients	Medicines	Medicine Prescriptions
Prescription ID	Patient ID	Patient ID	Medicine Brand	Frequency
Differential Diagnosis	Medicine Brand	Patient's Age Category	-	FrequencyDd
-	-	Patient's BMI Category	-	Timing
-	-	Gender	-	Duration
-	-	-	-	Amount

MODEL 2:
K-NEAREST NEIGHBOR

OUTPUT DATA: MEDICINE LIST


```
{
    "SURBEX T (KOMB Vit C 500 mg, Niasinamida 100 mg, Kalsium Pantotenat
20 mg,Vit. B1 15 mg, Vit. B2 10 mg,Vit. B6 5 mg, Vit. B12 4 mg) TAB": 124,
    "CETIRIZINE 10 MG Tablet": 78,
    "PARACETAMOL 500 MG Tablet": 77,
    "AMBROXOL 30 MG Tablet": 48,
    "Acetylcystein 200mg Tablet": 33,
    "SANADRYL DMP (Dekstrometorfan HBr 10 mg, Difenhidramin HCl 12,5 mg,
Ammonium Cl 100 mg, Na Sitrat 50 mg, Mentol 1 mg/ 5ML) SYRUP, BOTOL
120ML": 23,
    "DEXAMETHASONE 0.5 MG Tablet": 20,
    "PARACETAMOL 600 MG Tablet": 19,
    "CETIRIZINE 5 MG/5 ML Botol": 18,
    "ELKANA Syrup Botol 60ML": 18
}
```

MODEL 2:
K-NEAREST NEIGHBOR

OUTPUT DATA: MEDICINE DOSAGE

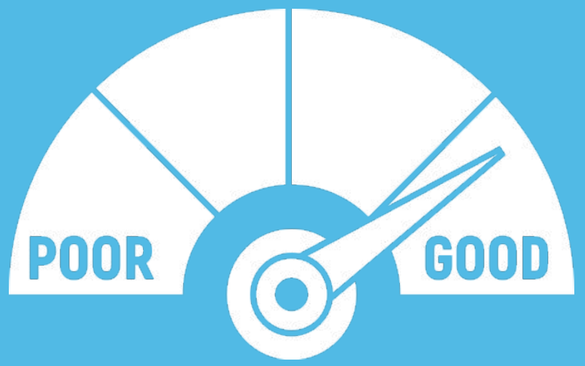

```
"SURBEX T (KOMB Vit C 500 mg, Niasinamida 100 mg, Kalsium Pantotenat 20
 mg,Vit. B1 15 mg, Vit. B2 10 mg,Vit. B6 5 mg, Vit. B12 4 mg) TAB": {
    "frequency": 1.0,
    "frequencyDd": 1.0,
    "timing": "Setelah Makan",
    "duration": 5.0,
    "amount": 5.0
}
```

EVALUATION: PERFORMANCE METRICS

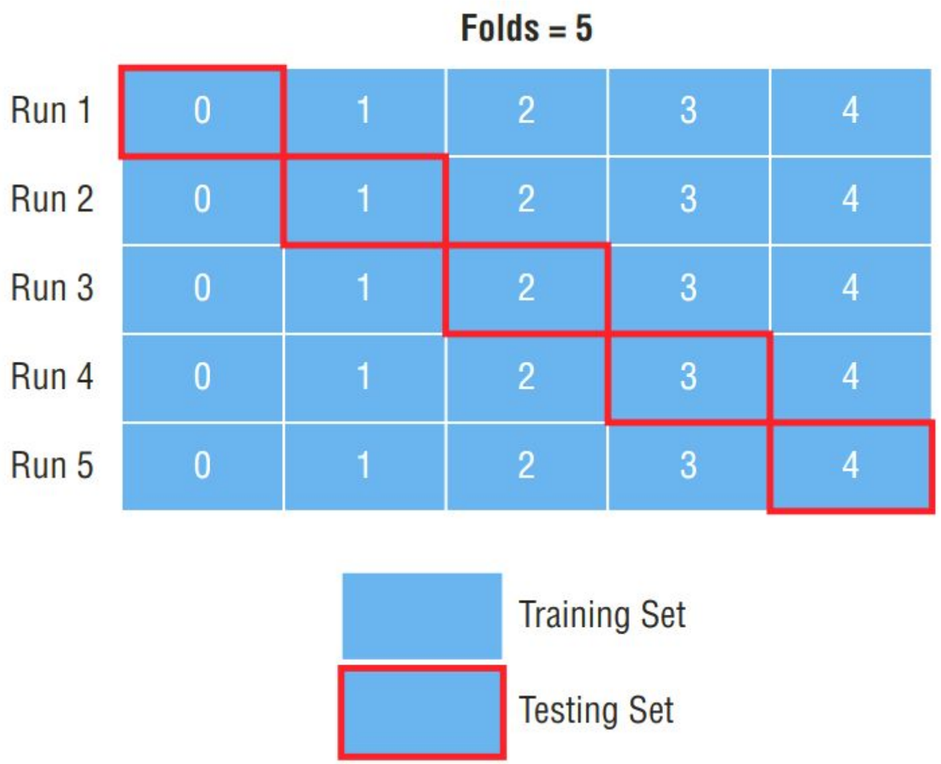
Recall@k:
The proportion of relevant medicines found in the recommendation medicines set

Precision@k:
Represents the proportion of recommended medicines (in the top k set; k=10) that are relevant

F1 Score:
Harmonic mean of both recall and precision



EVALUATION: K-FOLD CROSS VALIDATION



K Folds = 10
COLLABORATIVE FILTERING

Metric	Value
Recall@k	77.04
Precision@k	24.19
F1 Score	36.67
Runtime	2.14 s/item

K Folds = 10
K-NEAREST NEIGHBOR

Metric	Value
Recall@k	63.51
Precision@k	15.11
F1 Score	23.69
Runtime	1.06 s/item

MODEL REFINEMENT: BEST SCENARIOS

COLLABORATIVE FILTERING

Threshold = 2.01

Metric	Value
Recall@k	66.30
Precision@k	23.69
F1 Score	33.78
Runtime	1.48 s/item

KNN

K ICD = 151 , K Patient = 51

Metric	Value
Recall@k	63.91
Precision@k	15.20
F1 Score	23.84
Runtime	1.07 s/item

KNN

K Patient = 150, K ICD = 40

Threshold: Similarity = 0.04 , Med = 0.4

Metric	Value
Recall@k	62.20
Precision@k	14.73
F1 Score	23.12
Runtime	1.34 s/item

MODEL REFINEMENT: BEST SCENARIOS

COLLABORATIVE FILTERING

Threshold = 2.01

Metric	Value
Recall@k	66.30
Precision@k	23.69
F1 Score	33.78
Runtime	1.48 s/item

KNN

K ICD = 151 , K Patient = 51

Metric	Value
Recall@k	63.91
Precision@k	15.20
F1 Score	23.84
Runtime	1.07 s/item

KNN

K Patient = 150, K ICD = 40

Threshold: Similarity = 0.04 , Med = 0.4

Metric	Value
Recall@k	62.20
Precision@k	14.73
F1 Score	23.12
Runtime	1.34 s/item

MODEL DEPLOYMENT

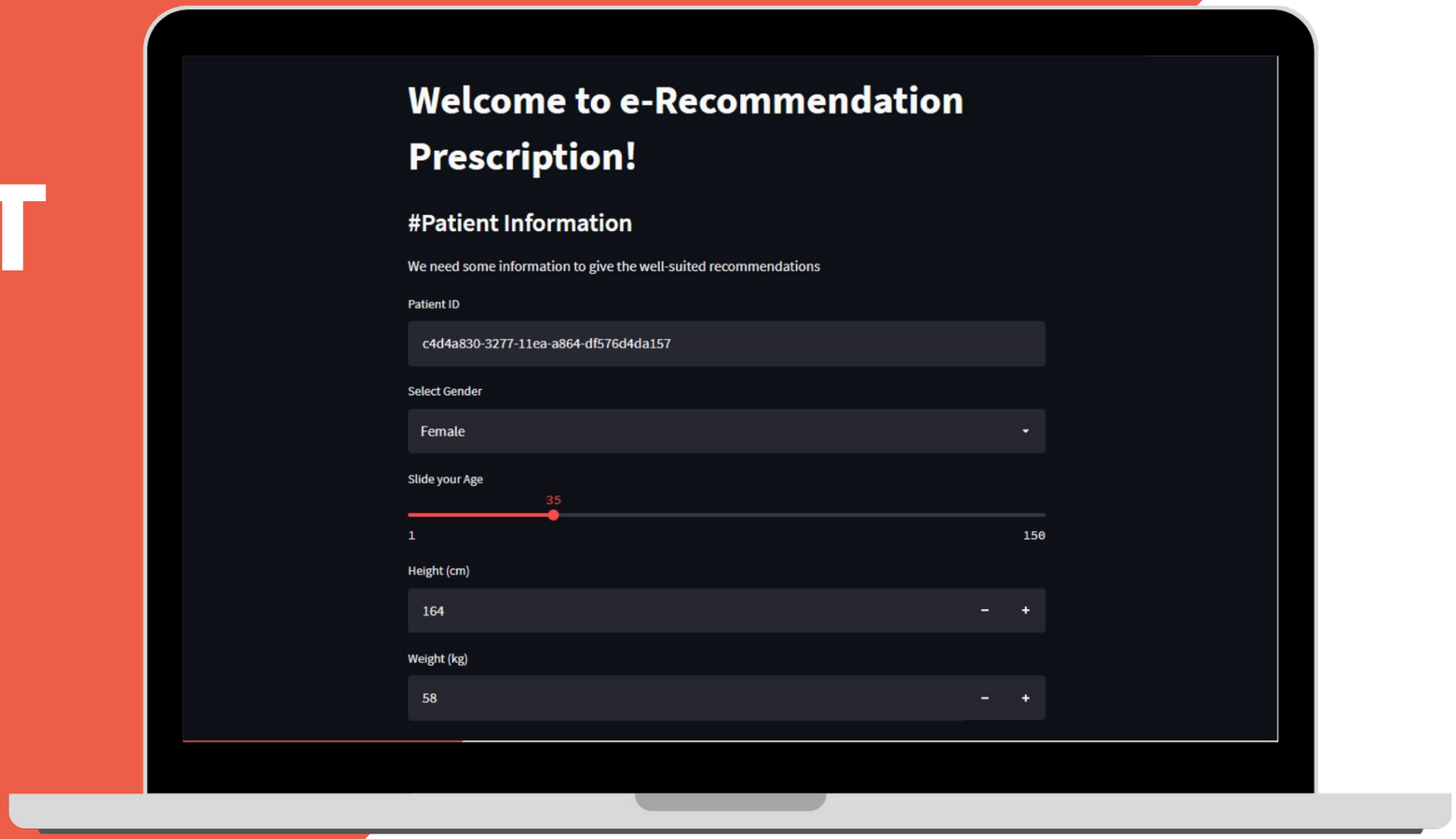
User Input:

Patient Information:

- Patient ID
- Gender
- Age
- Height (cm)
- Weight (kg)

Differential Diagnosis

- Differential Diagnosis 1- 5



MODEL DEPLOYMENT

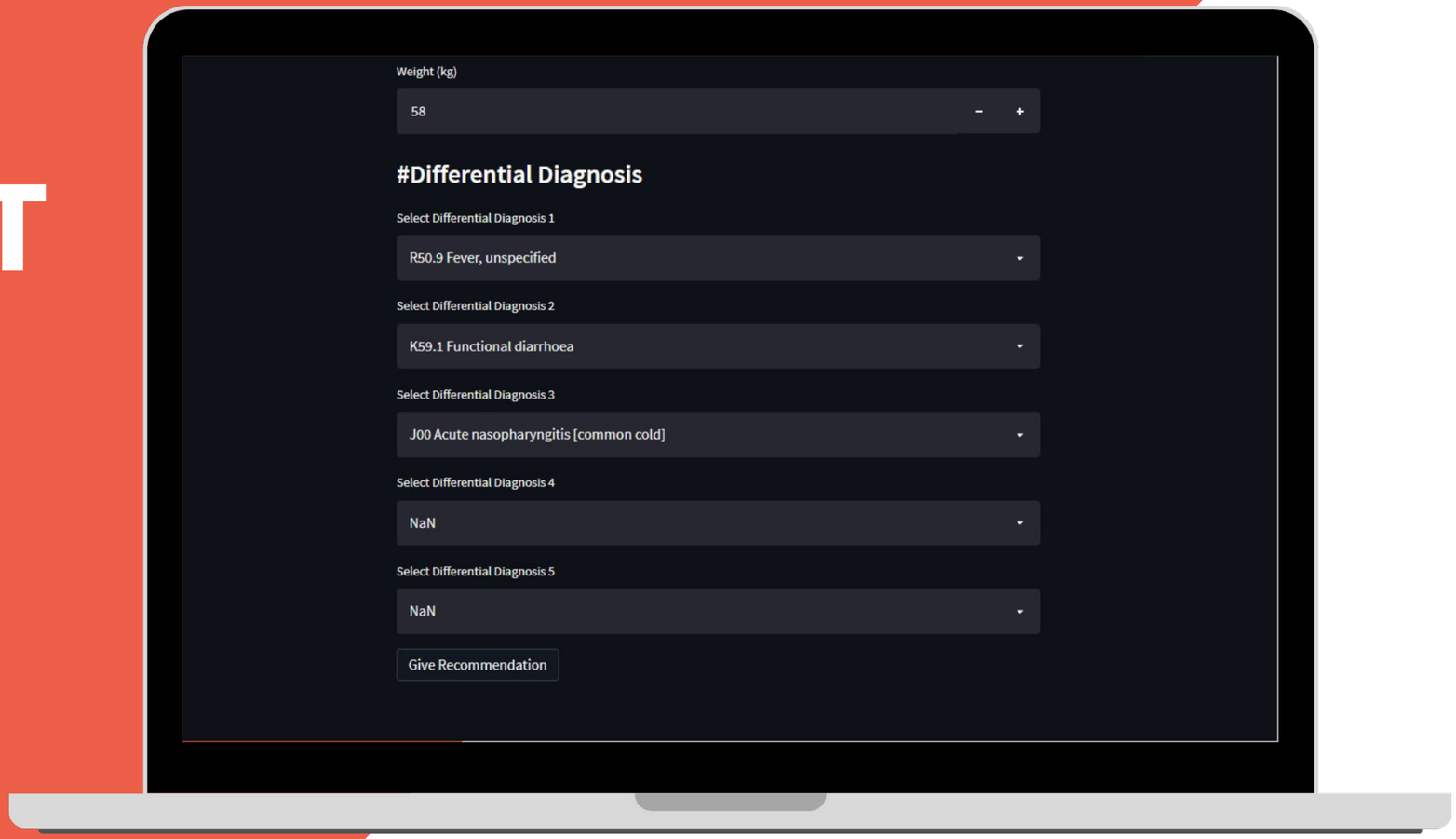
User Input:

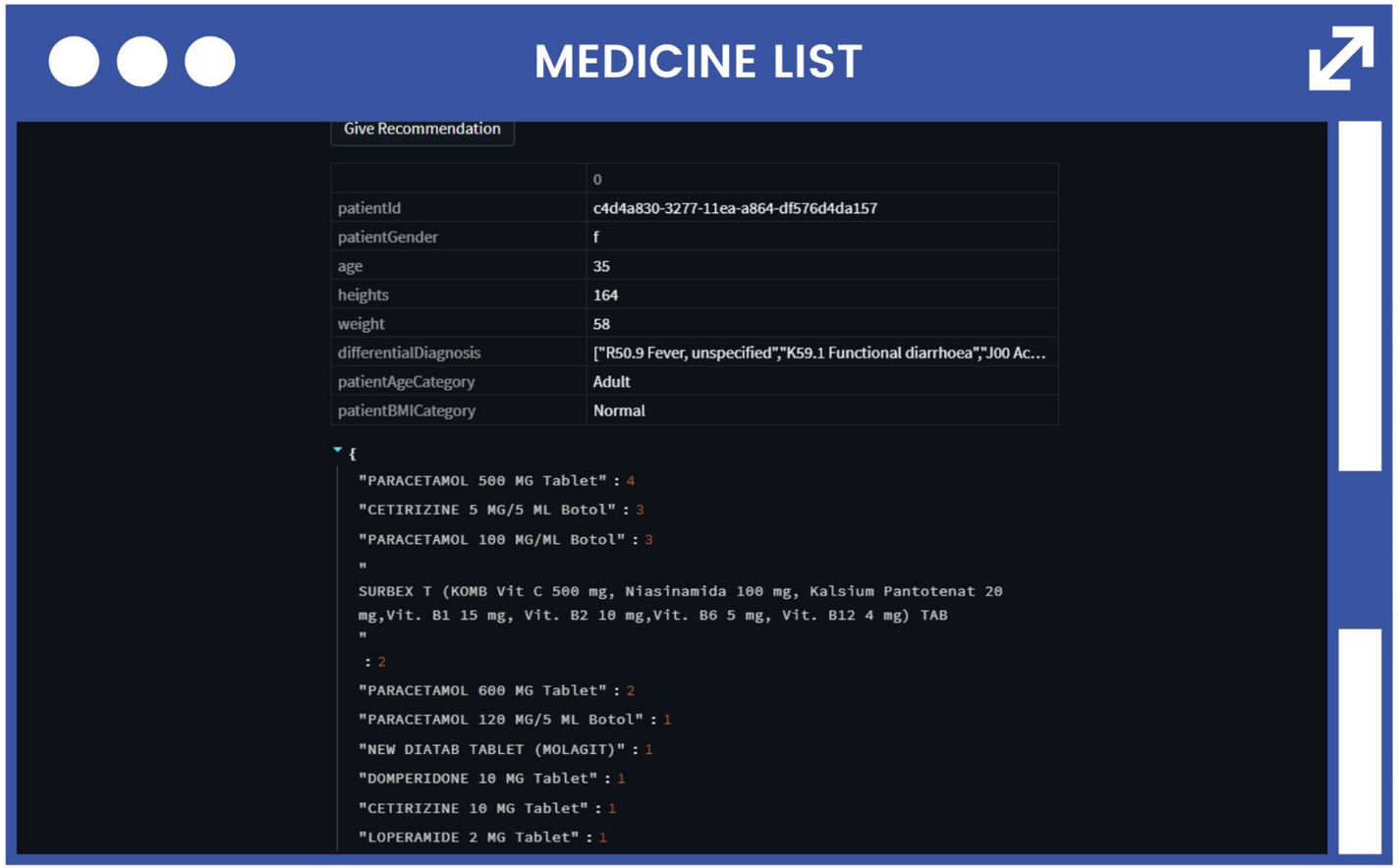
Patient Information:

- Patient ID
- Gender
- Age
- Height (cm)
- Weight (kg)

Differential Diagnosis

- Differential Diagnosis 1- 5





METHODOLOGY

MODEL DEPLOYMENT

OUTPUT

MEDICINE DOSAGE

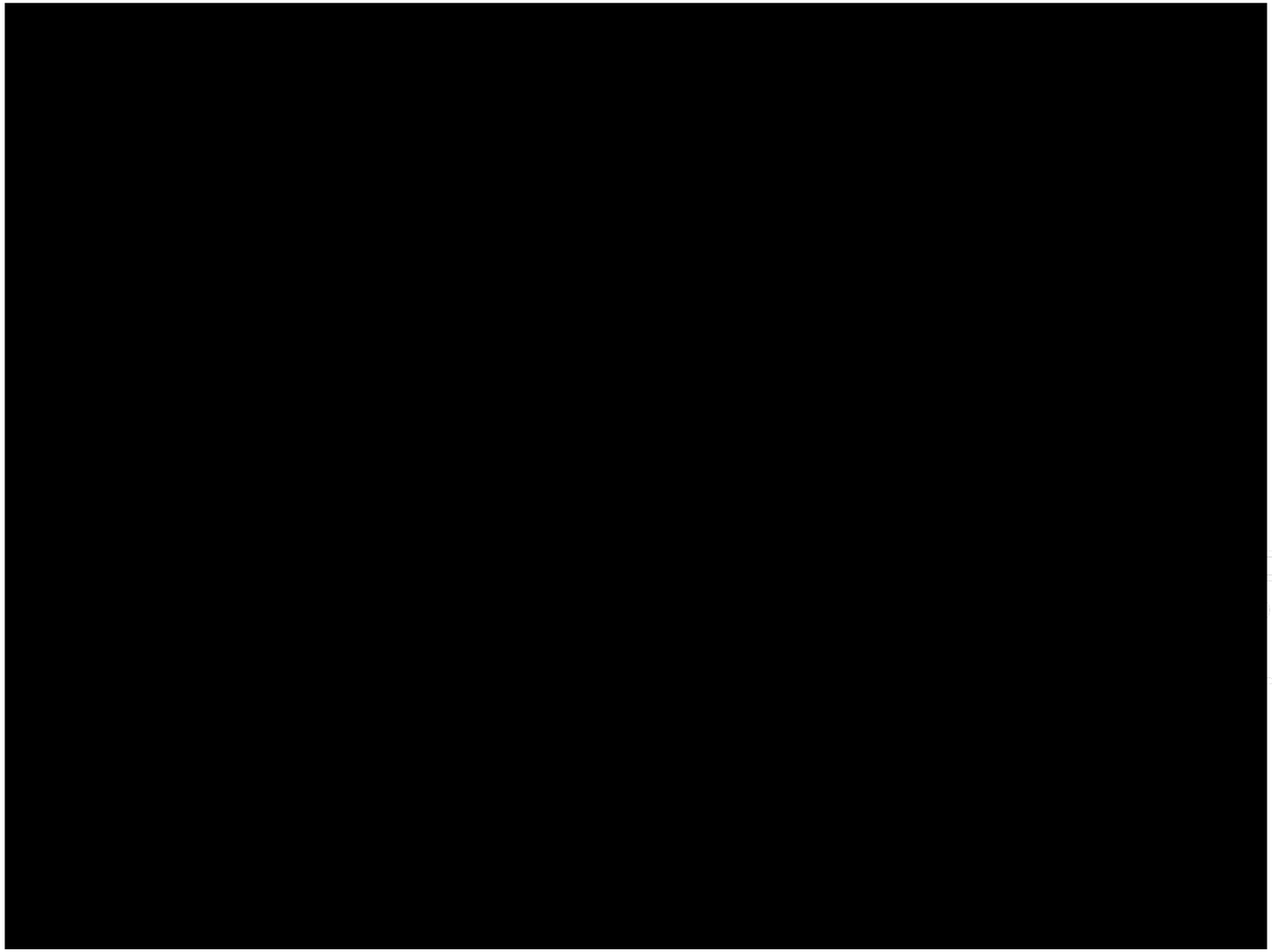

```
{
  "PARACETAMOL 500 MG Tablet" : {
    "frequency" : 3
    "frequencyDd" : 1
    "timing" : "Setelah Makan"
    "duration" : 3
    "amount" : 15
  }
  "CETIRIZINE 5 MG/5 ML Botol" : {
    "frequency" : 1
    "frequencyDd" : 5
    "timing" : "Setelah Makan"
    "duration" : 5
    "amount" : 1
  }
  "PARACETAMOL 100 MG/ML Botol" : {
    "frequency" : 3
    "frequencyDd" : 1
    "timing" : "Setelah Makan"
    "duration" : 5
    "amount" : 1
  }
  "SURBEX T (KOMB Vit C 500 mg, Niasinamida 100 mg, Kalsium Pantotenat 20
  mg,Vit. B1 15 mg, Vit. B2 10 mg,Vit. B6 5 mg, Vit. B12 4 mg) TAB" :
  {
```


METHODOLOGY

MODEL DEPLOYMENT

OUTPUT

MODEL DEPLOYMENT



CONCLUSION

Best model developed is collaborative filtering with threshold 2.01. This model produce recall@k of 66.30, precision@k 23.69, f1 score 33.78, and runtime 1.48 s/item.

THANK YOU!

dave.peerera@student.sgu.ac.id