CHAPTER 5
IMPLEMENTATION AND TESTING

5.1 Implementation

This section will explain how to use the code used in this program.

```
116. set totaldata :=
117. (SELECT COUNT(*) FROM tblTempData);
118.
119. set totalsatisfied :=
120. (SELECT COUNT(*) FROM tblTempData
121. WHERE classification = 'satisfied');
122.
123. set totaldissatisfied :=
124. (SELECT COUNT(*) FROM tblTempData
125. WHERE classification = 'neutral or dissatisfied');
```

Line 116 to 125 is used to count total data, total passenger satisfied, and total passenger dissatisfied.

```
464. set entrophyval :=
465.                                                    (-
  (totaldissatisfied/totaldata)*log2(totaldissatisfied/totaldata))
466.                              +                     (-
  (totalsatisfied/totaldata)*log2(totalsatisfied/totaldata));
```

Line 464 to 466 is used to calculate entrophy for total data.

```
471. INSERT INTO tblCount (informasi, jumlahdata, satisfied,
  dissatisfied)
472.        SELECT DISTINCT (a.inflight_wifi_service) AS 'INFLIGHT
  WIFI SERVICE',
473.        COUNT(a.inflight_wifi_service) AS JUMLAHDATA,
474.        (
475.          SELECT COUNT(*) FROM tblTempData AS b
476.          WHERE b.classification = 'satisfied' AND
477.          b.inflight_wifi_service = a.inflight_wifi_service
```

```
478.        ) AS satisfied,
479.        (
480.          SELECT COUNT(*) FROM tblTempData AS b
481.          WHERE b.classification = 'neutral or dissatisfied'
  AND
482.          b.inflight_wifi_service = a.inflight_wifi_service
483.        ) AS dissatisfied
484.        FROM tblTempData AS a
485.        GROUP BY a.inflight_wifi_service;
486.
487.        UPDATE tblCount set atribut = 'INFLIGHT WIFI SERVICE'
488.        WHERE atribut IS NULL;
489.
490.        INSERT INTO tblCount (informasi, jumlahdata, satisfied,
  dissatisfied)
491.        SELECT    DISTINCT    (a.departure_arrival_time)    AS
  'DEPARTURE ARRIVAL TIME',
492.        COUNT(a.departure_arrival_time) AS JUMLAHDATA,
493.        (
494.          SELECT COUNT(*) FROM tblTempData AS b
495.          WHERE b.classification = 'satisfied' AND
496.          b.departure_arrival_time = a.departure_arrival_time
497.        ) AS satisfied,
498.        (
499.          SELECT COUNT(*) FROM tblTempData AS b
500.          WHERE b.classification = 'neutral or dissatisfied'
  AND
501.          b.departure_arrival_time = a.departure_arrival_time
502.        ) AS dissatisfied
503.        FROM tblTempData AS a
      GROUP BY a.departure_arrival_time;
```

Line 471 to 503 is used to count total data, total satisfied, totaldissatisfied from a certain value.

```
737.     UPDATE tblCount set nilaiI =
```

```
738.            (-
   (satisfied/jumlahdata)*log2(satisfied/jumlahdata))
739.            +(-
   (dissatisfied/jumlahdata)*log2(dissatisfied/jumlahdata))
740.            WHERE satisfied <> 0 AND dissatisfied <> 0;
741.
742.            UPDATE tblCount set nilaiI = 0
743.            WHERE nilaiI IS NULL;
```

Line 737 to 743 is used to calculate entrophy of a certain attributes.

```
754.     INSERT INTO tblSementara(atribut, gain)
755.            SELECT atribut,
756.            entrophyval - SUM((jumlahdata/totaldata)*nilaiI)
757.            AS HITUNGGAIN
758.            FROM tblCount
759.            GROUP BY atribut;
760.
761.            UPDATE tblCount set gain =
762.            ROUND ((
763.              SELECT tblSementara.gain
764.              FROM tblSementara
765.              WHERE tblSementara.atribut = tblCount.atribut
766.            ), 4);
```

Line 754 to 766 is used to calculate all attributes gain.

5.2 Results

Some of the result from the executed program is shown below.

attribut	informasi	jumlahdata	satisfied	dissatisfied	nilail	Gain
totaldata	NULL	100	46	54	0.9954	0
inflight wifi service	0	1	1	0	0	0.1932
inflight wifi service	1	16	6	10	0.9544	0.1932
inflight wifi service	2	30	8	22	0.8366	0.1932
inflight wifi service	3	29	11	18	0.9576	0.1932

inflight wifi service	4	14	10	4	0.8631	0.1932
inflight wifi service	5	10	10	0	0	0.1932
departure arrival time	0	6	2	4	0.9183	0.0798
departure arrival time	1	14	11	3	0.7496	0.0798
departure arrival time	2	19	9	10	0.998	0.0798
departure arrival time	3	22	12	10	0.994	0.0798
departure arrival time	4	20	6	14	0.8813	0.0798
departure arrival time	5	19	6	13	0.8997	0.0798
easy of booking	0	3	1	2	0.9183	0.0387
easy of booking	1	15	7	8	0.9968	0.0387
easy of booking	2	26	8	18	0.8905	0.0387
easy of booking	3	32	15	17	0.9972	0.0387
easy of booking	4	13	8	5	0.9612	0.0387
easy of booking	5	11	7	4	0.9457	0.0387
gate location	1	18	10	8	0.9911	0.0242
gate location	2	22	12	10	0.994	0.0242
gate location	3	31	14	17	0.9932	0.0242
gate location	4	17	5	12	0.874	0.0242
gate location	5	12	5	7	0.9799	0.0242
food and drink	1	10	2	8	0.7219	0.0725
food and drink	2	28	10	18	0.9403	0.0725
food and drink	3	16	11	5	0.896	0.0725
food and drink	4	25	10	15	0.971	0.0725
food and drink	5	21	13	8	0.9587	0.0725
online boarding	0	1	0	1	0	0.2231
online boarding	1	9	3	6	0.9183	0.2231
online boarding	2	17	3	14	0.6723	0.2231
online boarding	3	18	2	16	0.5033	0.2231
online boarding	4	34	22	12	0.9367	0.2231
online boarding	5	21	16	5	0.7919	0.2231
seat comfort	1	6	1	5	0.65	0.0624
seat comfort	2	17	5	12	0.874	0.0624
seat comfort	3	12	4	8	0.9183	0.0624
seat comfort	4	32	20	12	0.9544	0.0624
seat comfort	5	33	16	17	0.9993	0.0624

inflight entertainment	1	9	1	8	0.5033	0.1158
inflight entertainment	2	22	5	17	0.7732	0.1158
inflight entertainment	3	8	3	5	0.9544	0.1158
inflight entertainment	4	28	18	10	0.9403	0.1158
inflight entertainment	5	33	19	14	0.9834	0.1158
on board service	1	10	3	7	0.8813	0.0724
on board service	2	12	3	9	0.8113	0.0724
on board service	3	16	4	12	0.8113	0.0724
on board service	4	32	18	14	0.9887	0.0724
on board service	5	30	18	12	0.971	0.0724
leg room service	1	9	2	7	0.7642	0.228
leg room service	2	23	4	19	0.6666	0.228
leg room service	3	21	5	16	0.7919	0.228
leg room service	4	22	15	7	0.9024	0.228
leg room service	5	25	20	5	0.7219	0.228
baggage handling	1	7	4	3	0.9852	0.0794
baggage handling	2	9	2	7	0.7642	0.0794
baggage handling	3	25	6	19	0.795	0.0794
baggage handling	4	27	15	12	0.9911	0.0794
baggage handling	5	32	19	13	0.9745	0.0794
checking service	1	15	7	8	0.9968	0.0437
checking service	2	8	4	4	1	0.0437
checking service	3	25	12	13	0.9988	0.0437
checking service	4	30	9	21	0.8813	0.0437
checking service	5	22	14	8	0.9457	0.0437
inflight service	1	4	2	2	1	0.0502
inflight service	2	15	5	10	0.9183	0.0502
inflight service	3	14	3	11	0.7496	0.0502
inflight service	4	31	15	16	0.9992	0.0502
inflight service	5	36	21	15	0.9799	0.0502
cleanliness	1	7	0	7	0	0.1143
cleanliness	2	20	5	15	0.8113	0.1143
cleanliness	3	18	10	8	0.9911	0.1143
cleanliness	4	27	14	13	0.999	0.1143
cleanliness	5	28	17	11	0.9666	0.1143

Table 2 : Iteration 1 from 100 data

From the table above the data will be calculated again using the same formula but with only the data from the best gain and using one of the value. In this case it will be leg room service with value 1. The table below is the calculation after using the data with only the best gain.

attribut	informasi	jumlahdata	satisfied	dissatisfied	nilail	Gain
totaldata	NULL	9	2	7	0.7642	0
inflight wifi service	0	1	1	0	0	0.7642
inflight wifi service	1	4	0	4	0	0.7642
inflight wifi service	2	2	0	2	0	0.7642
inflight wifi service	3	1	0	1	0	0.7642
inflight wifi service	4	1	1	0	0	0.7642
departure arrival time	0	3	1	2	0.9183	0.4581
departure arrival time	1	1	1	0	0	0.4581
departure arrival time	3	2	0	2	0	0.4581
departure arrival time	4	1	0	1	0	0.4581
departure arrival time	5	2	0	2	0	0.4581
easy of booking	0	2	1	1	1	0.1814
easy of booking	1	4	1	3	0.8113	0.1814
easy of booking	2	1	0	1	0	0.1814
easy of booking	3	2	0	2	0	0.1814
gate location	1	3	1	2	0.9183	0.152
gate location	2	3	1	2	0.9183	0.152
gate location	3	2	0	2	0	0.152
gate location	4	1	0	1	0	0.152
food and drink	1	2	0	2	0	0.542
food and drink	2	4	0	4	0	0.542
food and drink	3	1	1	0	0	0.542
food and drink	4	2	1	1	1	0.542
online boarding	1	1	0	1	0	0.4581
online boarding	2	2	0	2	0	0.4581
online boarding	3	2	0	2	0	0.4581
online boarding	4	3	2	1	0.9183	0.4581
online boarding	5	1	0	1	0	0.4581
seat comfort	2	3	0	3	0	0.4581
seat comfort	3	1	0	1	0	0.4581
seat comfort	4	3	2	1	0.9183	0.4581

seat comfort	5	2	0	2	0	0.4581
inflight entertainment	1	1	1	0	0	0.4581
inflight entertainment	2	3	0	3	0	0.4581
inflight entertainment	3	1	0	1	0	0.4581
inflight entertainment	4	3	1	2	0.9183	0.4581
inflight entertainment	5	1	0	1	0	0.4581
on board service	1	2	1	1	1	0.3198
on board service	2	1	0	1	0	0.3198
on board service	3	2	1	1	1	0.3198
on board service	4	3	0	3	0	0.3198
on board service	5	1	0	1	0	0.3198
leg room service	1	9	2	7	0.7642	0
baggage handling	1	1	1	0	0	0.4036
baggage handling	3	4	1	3	0.8113	0.4036
baggage handling	4	2	0	2	0	0.4036
baggage handling	5	2	0	2	0	0.4036
checking service	2	2	1	1	1	0.3198
checking service	3	2	1	1	1	0.3198
checking service	4	4	0	4	0	0.3198
checking service	5	1	0	1	0	0.3198
inflight service	1	1	1	0	0	0.542
inflight service	2	2	0	2	0	0.542
inflight service	3	1	0	1	0	0.542
inflight service	4	3	0	3	0	0.542
inflight service	5	2	1	1	1	0.542
cleanliness	1	1	0	1	0	0.4581
cleanliness	2	3	0	3	0	0.4581
cleanliness	3	1	0	1	0	0.4581
cleanliness	4	3	2	1	0.9183	0.4581
cleanliness	5	1	0	1	0	0.4581

Table 3 : Iteration 2 from 100 data

From the table above the best gain is inflight wifi service. After taking one of the value that from inflight wifi service that is 0 we can see the number of entrophy is 0 so one of them should be the decision from this branch and that is satisfied because it has more than 1 value. With that this branch has ended with “leg room service” value 1 > “inflight wifi service” value

0 > class satisfied. After reaching a decision the data will become full data again but the calculation will be taking another branch that has not been visited. This process will repeat until there is no more branch can be calculated.

no	id	real	result
1	103774	nod	nod
2	103775	nod	nod
3	103776	satisfied	nod
4	103777	satisfied	satisfied
5	103778	satisfied	satisfied
6	103779	nod	nod
7	103780	nod	nod
8	103781	nod	satisfied
9	103782	satisfied	satisfied
10	103783	satisfied	satisfied
11	103784	nod	satisfied
12	103785	nod	nod
13	103786	nod	nod
14	103787	satisfied	nod
15	103789	nod	nod
16	103790	satisfied	satisfied
17	103791	nod	nod
18	103792	nod	nod
19	103793	satisfied	satisfied
20	103794	satisfied	satisfied

Table 4 : Comparasion between real class and decision tree result from 100 train data

no	id	real	result
1	103799	satisfied	satisfied
2	103800	nod	nod
3	103801	satisfied	nod
4	103802	nod	nod
5	103803	nod	nod
6	103804	nod	nod
7	103805	nod	nod
8	103806	nod	nod
9	103807	nod	nod
10	103808	satisfied	satisfied
11	103809	satisfied	satisfied
12	103811	satisfied	satisfied
13	103813	nod	nod
14	103816	satisfied	nod
15	103819	nod	nod
16	103820	satisfied	satisfied
17	103821	nod	nod
18	103823	satisfied	satisfied
19	103824	nod	nod
20	103825	satisfied	satisfied

Table 5 : Comparasion between real class and decision tree result from 200 train data

From the tables above that is acquired from this research 4.5 algorithm has around 80-90% accuracy. The data used for the test is only good data. The good data in this research is the one that have a result when calculated using decision tree from the respective amount of train data.

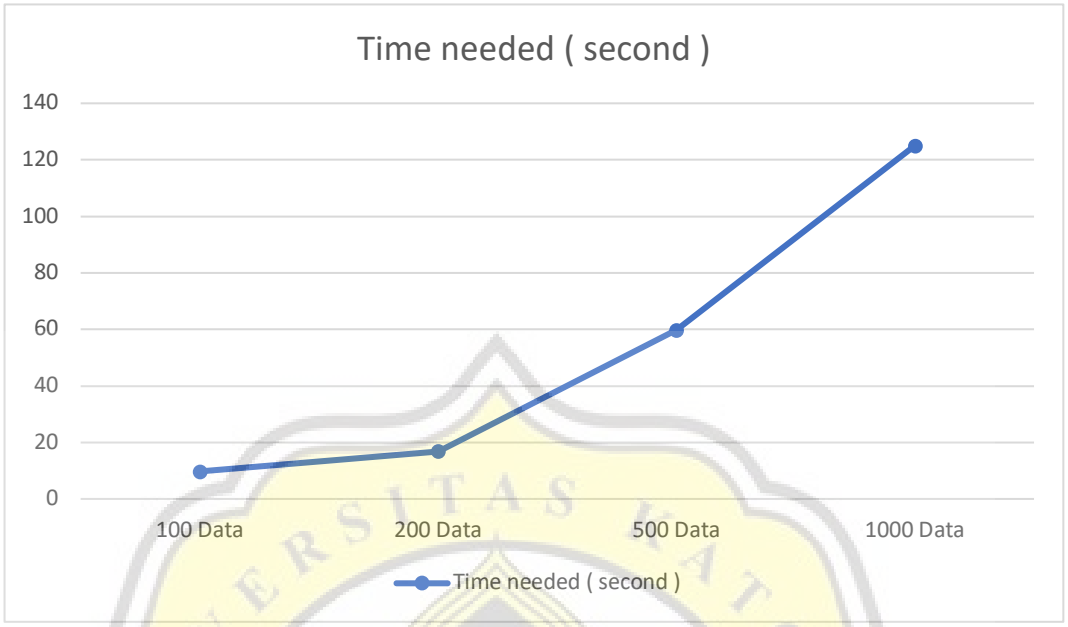

Figure 3 : Difference time needed for X amount of data in seconds