

BAB IV HASIL DAN PEMBAHASAN

4.1. Hasil Pengumpulan Data

Pada bab ini peneliti akan menjelaskan hasil penelitian yang sedang dikerjakan mengenai pengimplementasian data mining untuk memprediksi perkembangan belajar santri menggunakan metode *K-Nearest Neighbor* dengan mengimplementasikan data mining. Peneliti melakukan observasi terhadap objek penelitian dan wawancara kepada koordinator lembaga untuk mengumpulkan data yang dibutuhkan diantaranya adalah data nilai Tajwid, BTQ, Furudul 'Ainiyah, Fiqih dan Aqidah. Adapun hasil dari pengumpulan data tersebut adalah sebagai berikut:

4.1.1 Hasil Observasi

Observasi terhadap objek dilakukan pada bulan Maret di lembaga tamhidiyah wilayah al-hasyimiyah dengan tujuan untuk memberikan gambaran nyata tentang kegiatan belajar mengajar serta kegiatan keseharian santri lembaga tamhidiyah itu sendiri. Darn wawancara kepada coordinator lembaga dilakukan untuk mengumpulkan beberapa data yang dibutuhkan.

4.1.2. Hasil Wawancara

Wawancara dilakukan secara langsung kepada koordinator lembaga yaitu ustadzah Laily Sholihah di kantor lembaga tamhidiyah pada awal bulan April. Adapun hasil wawancara adalah sebagai berikut :

- a. Proses belajar mengajar di lembaga tamhidiyah ini dilakukan pada malam hari yang ditemani oleh seorang guru dalam setiap kelas.
- b. Santri akan dinyatakan lulus dari lembaga jika santri tersebut telah tuntas dalam menguasai semua materi yang telah ditentukan dan nilai akhir mereka sudah memenuhi syarat kelulusan.

Dari hasil penelitian yang dilakukan dengan observasi dan wawancara maka didapat data dari bagian KA TU lembaga yang bersangkutan dengan jumlah data ialah 231. dengan variable data yang digunakan adalah data nilai Tajwid, BTQ, Furudul ‘Ainiyah, Fiqih dan Aqidah Dari 231 data santri yang diambil ialah data santri pada angkatan tahun 2018 sampai 2020. Adapun jumlah data santri angkatan 2018 adalah 85 santri adapun santri yang berstatus Baru adalah 49 dan yang berstatus Lama adalah 36. Untuk santri angkatan 2019 berjumlah 78 santri adapun santri yang berstatus baru adalah 51 santri dan yang berstatus lama adalah 27 santri. Sedangkan jumlah pada data santri angkatan 2020 adalah 66 santri, adapun jumlah santri yng berstatus baru adalah 45 santri, sedangkan yang berstatus lama adalah 21 santri. Jadi, dari 231 santri jumlah keseluruhan santri, santri yang berstatus baru ada 145 dan jumlah santri yang berstatus lama ada 86 santri. Dari hasil penelitian yang dilakukan dengan observasi dan wawancara maka didapat dataset sebagai berikut :

Tabel 4.1 Dataset

NO	NAMA	Status	Angkatan	Tajwid	BTQ	FA	Fiqih	Aqidah	UJIAN	KET
1	Umrotul Hasanah	Lama	2018	87	85	80	77	82	82,2	Lulus
2	Indah Indrawati	Baru	2018	63	70	54	64	65	63,2	Tidak
3	Tartila	Baru	2018	95	74	85	63	91	81,6	Lulus
4	Nur Mardatillah	Baru	2018	85	70	81	77	80	78,6	Lulus
5	Nur Ayu Febriana	Baru	2018	71	75	64	53	76	67,8	Tidak
6	Aliffida Ussi Adalia	Baru	2018	70	70	68	70	68	69,2	Tidak
7	Siti Hamidah	Baru	2018	83	95	85	70	80	82,6	Lulus
8	Siti Fatimah	Baru	2018	79	63	63	70	54	65,8	Tidak
9	Ummi Rifqotur Rofiqoh	Baru	2018	81	86	75	73	84	79,8	Lulus
10	Zakiyatus Safiroh	Baru	2018	86	76	85	70	75	78,4	Lulus
11	Nadzirotul Jamilah	Baru	2018	86	85	76	83	76	81,2	Lulus
12	Nur Hilla Lina Kamila	Baru	2018	71	69	74	70	73	71,4	Lulus
13	Isyti Rozanah	Lama	2018	76	73	91	85	72	79,4	Lulus
14	Mahbubatul Hafifi	Lama	2018	85	74	82	86	72	79,8	Lulus
15	Hanifatur Rodiyah	Lama	2018	80	80	73	95	75	80,6	Lulus
16	Ika Maziyyatus Sholihah	Lama	2018	75	75	81	72	70	74,6	Lulus
17	Nur Khofifah	Baru	2018	63	69	61	71	68	66,4	Tidak
18	Wildatul Aluf	Lama	2018	80	85	73	74	71	76,6	Lulus
19	Yeni Cahyani	Baru	2018	54	65	66	40	59	56,8	Tidak
20	Siti Qomariah	Baru	2018	74	90	75	75	70	76,8	Lulus

Lanjutan Dataset

21	Kholifatun Nisa'	Lama	2018	82	87	90	84	81	84,8	Lulus
22	Kamaliyah	Lama	2018	78	76	72	70	81	75,4	Lulus
23	Nur Hasanah	Baru	2018	78	68	71	77	70	72,8	Lulus
24	Nur Aini	Baru	2018	75	70	70	69	72	71,2	Lulus
25	Aprilia Dwi K	Baru	2018	62	55	57	47	68	57,8	Tidak
26	Nofi Ferandita	Lama	2018	76	90	80	68	75	77,8	Lulus
27	Khoirun Nisa'	Baru	2018	77	83	85	81	73	79,8	Lulus
28	Khorul Bariroh	Baru	2018	85	82	70	76	85	79,6	Lulus
29	Siti Romlah	Lama	2018	85	50	64	50	90	67,8	Tidak
30	Yuliana Putri	Baru	2018	60	59	68	56	59	60,4	Tidak
31	Lailatul Fadila	Baru	2018	54	50	63	58	69	58,8	Tidak
32	Ina Susanti	Baru	2018	63	58	43	64	60	57,6	Tidak
33	Tiara Ainur Rizki	Lama	2018	82	87	76	84	75	80,8	Lulus
34	Faridatul Aini	Lama	2018	70	73	71	72	76	72,4	Lulus
35	Dwi Putri Lianis Kurli	Lama	2018	75	79	90	85	80	81,8	Lulus
36	Dwiki Oktafiana W	Lama	2018	75	78	81	79	86	79,8	Lulus
37	Aisyah	Lama	2018	86	94	85	77	72	82,8	Lulus
38	Dewi Setya R N	Lama	2018	63	58	62	60	73	63,2	Tidak
39	Halimatus Sa'diyah	Baru	2018	71	70	63	71	68	68,6	Tidak
40	Nuril Ummu Latifah	Baru	2018	60	55	60	68	65	61,6	Tidak
41	Inni Qomariah	Baru	2018	65	75	63	61	62	65,2	Tidak
42	Gina Aulia	Baru	2018	50	59	75	60	70	62,8	Tidak
43	Afinatul Qodoriyah	Baru	2018	55	60	68	57	63	60,6	Tidak
44	Jelang Januari Fitri	Lama	2018	80	83	95	78	83	83,8	Lulus
45	Miftahis sa'adah Uliya	Lama	2018	60	75	85	79	80	75,8	Lulus
46	Zakia Fitri Amalia	Baru	2018	76	82	80	83	78	79,8	Lulus
47	Zahrotul Aini	Lama	2018	85	87	83	85	96	87,2	Lulus
48	Latifatul Ma'rifah N.H	Lama	2018	83	87	79	93	85	85,4	Lulus
49	Siti Rohaniyah	Baru	2018	76	70	68	72	76	72,4	Lulus
50	Diana Nur Faizah	Baru	2018	72	75	80	75	79	76,2	Lulus
51	Rafika Milania	Baru	2018	85	91	89	90	87	88,4	Lulus
52	Siti Rosyidah	Baru	2018	76	85	80	73	75	77,8	Lulus
53	Ayu Midyah putri	Baru	2018	67	70	58	76	80	70,2	Lulus
54	Faizatul Rohmah	Baru	2018	55	63	70	69	72	65,8	Tidak
55	Dini Melinda Ulfatul A	Lama	2018	82	85	72	81	87	81,4	Lulus
56	Nafilatul Azizah	Baru	2018	71	65	75	85	70	73,2	Lulus
57	Muhtamilatul Rohmah	Lama	2018	80	95	80	70	58	76,6	Lulus
58	Ulfatun Nazilah	Lama	2018	80	55	73	90	90	77,6	Lulus
59	Nadiya Rosmila	Baru	2018	70	85	90	67	90	80,4	Lulus
60	Nur Laili	Baru	2018	75	90	82	86	80	82,6	Lulus
61	Devi Novita	Baru	2018	80	70	58	70	75	70,6	Lulus
62	Sri Ayu Putri N	Lama	2018	65	80	90	50	93	75,6	Lulus
63	Siti Aisyah	Baru	2018	80	80	90	87	70	81,4	Lulus
64	Viqi Fadilatul R	Lama	2018	75	57	76	73	85	73,2	Lulus
65	Indana Fikrotul Arifin	Lama	2018	76	65	75	90	70	75,2	Lulus
66	Zilfiyah Firdausy	Lama	2018	50	95	93	70	70	75,6	Lulus
67	Anis Fitria	Lama	2018	75	88	85	85	85	83,6	Lulus
67	Anis Fitria	Lama	2018	75	88	85	85	85	83,6	Lulus
68	Dzurrotul Muniroh A	Lama	2018	82	80	80	70	60	74,4	Lulus
69	Hanifatul Qomariyah	Baru	2018	95	85	89	70	70	81,8	Lulus
70	Zaimatul Umami	Baru	2018	78	75	87	85	83	81,6	Lulus

Lanjutan Dataset

71	Sofiatul Qudsiyah	Baru	2018	70	64	78	70	62	68,8	Tidak
72	Ifa Krismoning Tiyas	Baru	2018	60	71	80	60	68	67,8	Tidak
73	Siti Ibra Sonia	Lama	2018	75	85	82	73	87	80,4	Lulus
74	Dzurrotul Mashunah	Lama	2018	82	78	89	85	80	82,8	Lulus
75	Rindi Wulandari	Baru	2018	53	61	66	56	68	60,8	Tidak
76	Kamelia Yulia Nofita	Lama	2018	76	75	81	76	70	75,6	Lulus
77	Siti Fatimah Daniatul Fitriyah	Baru	2018	81	86	83	84	79	82,6	Lulus
78	Maulidatul Hasanah	Lama	2018	67	56	71	63	57	62,8	Tidak
79	Wardatul Munawwaroh	Lama	2018	91	85	70	80	80	81,2	Lulus
80	Elmiyatus Sholihah	Lama	2018	75	85	70	81	85	79,2	Lulus
81	Siti Kholilatul Jannah	Baru	2018	83	89	80	87	80	83,8	Lulus
82	Emilia Safitri	Baru	2018	76	84	80	87	85	82,4	Lulus
83	Rifki Antika	Lama	2018	65	60	78	80	80	72,6	Lulus
84	Ummi Nur Aisyah	Baru	2018	59	85	78	80	80	76,4	Lulus
85	Wahidah	Lama	2018	84	71	80	90	83	81,6	Lulus
86	Putri fadhillah	Baru	2019	72	69	73	85	77	75,2	Lulus
87	Yani Cantika T	Baru	2019	57	64	61	76	58	63,2	Tidak
88	Azizatul Layali	Baru	2019	60	52	64	55	61	58,4	Tidak
89	Ismi	Baru	2019	75	75	87	79	82	79,6	Lulus
90	Tsania Ikrimah	Baru	2019	85	70	80	75	86	79,2	Lulus
91	Putri Azizah	Baru	2019	90	80	92	80	80	84,4	Lulus
92	Nurul Jannah Safariah	Baru	2019	85	73	80	85	75	79,6	Lulus
93	Arika Ramadhany	Baru	2019	75	80	71	85	78	77,8	Lulus
94	Chindy Putri Ramadhany	Baru	2019	70	70	68	70	66	68,8	Tidak
95	Zulfatun Nazilah	Baru	2019	76	65	70	83	78	74,4	Lulus
96	Putri Mar'atus Sholehah	Baru	2019	79	70	65	84	73	74,2	Lulus
97	Faiqotul Izzah	Baru	2019	79	86	81	90	82	83,6	Lulus
98	Maslikha	Baru	2019	70	82	80	75	85	78,4	Lulus
99	Khalisatun Nuriyah	Baru	2019	80	70	95	94	70	81,8	Lulus
100	Siti Kholimatus Sholehah	Baru	2019	80	74	80	80	82	79,2	Lulus
101	Siti Azimul Faiqe	Lama	2019	65	85	78	75	85	77,6	Lulus
102	Inayatul Hidayah	Lama	2019	85	80	92	96	88	88,2	Lulus
103	Ahnas Finafisatil Aisyi	Lama	2019	81	78	83	80	85	81,4	Lulus
104	Nurul Ismah	Lama	2019	82	79	80	85	83	81,8	Lulus
105	Faiqotul Hilyah	Lama	2019	90	86	80	90	80	85,2	Lulus
106	Nurhidayah	Lama	2019	87	89	78	90	80	84,8	Lulus
107	Malihatul lu'lu'a	Baru	2019	70	74	78	65	85	74,4	Lulus
108	Qurrotul Aini	Baru	2019	50	65	45	70	76	61,2	Tidak
109	Alfiatul Qomariah	Baru	2019	81	85	76	85	80	81,4	Lulus
110	Fiqih Humaidah	Lama	2019	65	75	93	65	60	71,6	Lulus
111	Irna Karisma D	Baru	2019	80	63	82	85	73	76,6	Lulus
112	Nur Diana Rizqiyatin	Lama	2019	85	75	80	85	81	81,2	Lulus
113	Laili Sholihah	Baru	2019	83	72	81	79	87	80,4	Lulus
114	Siti Humairoh	Baru	2019	79	75	85	85	64	77,6	Lulus
115	Hulliyatus Saniah	Lama	2019	91	80	71	87	85	82,8	Lulus
116	Qudsiyatut Diana	Lama	2019	75	90	84	72	91	82,4	Lulus
117	Uswatun Hasanah	Lama	2019	80	85	75	81	80	80,2	Lulus
118	Amelia Putri	Lama	2019	67	80	85	85	70	77,4	Lulus
119	Sholehatus Mahdiah	Lama	2019	85	70	75	67	55	70,4	Lulus

Lanjutan Dataset

120	Siti Aisyah	Lama	2019	85	72	90	83	77	81,4	Lulus
121	Fitri Nur Syamsiah	Baru	2019	60	54	60	63	72	61,8	Tidak
122	Vini Ramadhaniah Sami	Baru	2019	50	85	80	95	58	73,6	Lulus
123	Anisa Qomariah	Baru	2019	66	64	56	65	40	58,2	Tidak
124	Kamaliatus Syamsiah	Lama	2019	84	87	75	88	73	81,4	Lulus
125	Fatin Fatiqoh	Baru	2019	86	75	84	75	81	80,2	Lulus
126	Rafika	Baru	2019	80	88	70	90	80	81,6	Lulus
127	Apriyanti	Baru	2019	90	65	80	78	75	77,6	Lulus
128	Diana Izza	Baru	2019	85	73	75	84	90	81,4	Lulus
129	Ning Masruroh	Lama	2019	62	75	95	71	65	73,6	Lulus
130	Jamilatul Mahyah	Baru	2019	85	73	85	70	74	77,4	Lulus
131	Nurul Faiqoh	Baru	2019	75	80	73	70	85	76,6	Lulus
132	Mirayanti	Baru	2019	70	73	85	80	69	75,4	Lulus
133	Nanik Mustika Agustin	Baru	2019	67	70	75	65	80	71,4	Lulus
134	Dinal Islamiyah	Lama	2019	90	84	70	85	60	77,8	Lulus
135	Fatimatus Zahro	Lama	2019	77	65	73	69	70	70,8	Lulus
136	Nadia Latifa	Baru	2019	85	80	73	85	75	79,6	Lulus
137	Arfa AnNisa' Awwalina Maulita	Baru	2019	85	90	67	71	75	77,6	Lulus
138	Yulianti	Baru	2019	75	81	77	50	56	67,8	Tidak
139	Syifa Fauziyah Nurul Izzah	Baru	2019	80	95	85	50	89	79,8	Lulus
140	Anisa Rifdah Imami	Lama	2019	80	72	68	70	58	69,6	Tidak
141	Hilyatus Syafiah	Lama	2019	85	50	53	60	71	63,8	Tidak
142	Diana Maufiroh	Lama	2019	60	75	85	50	53	64,6	Tidak
143	Icuk Putri Utami	Baru	2019	60	90	65	74	90	75,8	Lulus
144	Lailatul Fitriah	Lama	2019	85	95	85	70	67	80,4	Lulus
145	Nida Uswatun Hasanah	Lama	2019	80	65	81	73	80	75,8	Lulus
146	Silvy Qotrun Nada	Lama	2019	82	77	85	70	80	78,8	Lulus
147	Zahro Anisa	Baru	2019	88	83	65	70	63	73,8	Lulus
148	Sa'diyatur Rohma	Baru	2019	70	80	75	80	71	75,2	Lulus
149	Amelatus Zahro	Baru	2019	73	64	75	84	77	74,6	Lulus
150	Ira Alfaida	Baru	2019	85	78	80	76	83	80,4	Lulus
151	Alfin Nurir Rahman	Lama	2019	72	88	85	70	87	80,4	Lulus
152	Nurul Muthmainnah	Baru	2019	73	76	80	82	80	78,2	Lulus
153	Makkiyatun Nufus	Baru	2019	65	95	55	73	64	70,4	Lulus
154	Lailatul Mukarromah	Baru	2019	73	75	80	80	85	78,6	Lulus
155	Maulidatun Nafizatus Zahro	Baru	2019	70	76	80	74	83	76,6	Lulus
156	Cindy Nur Diana Alifia	Baru	2019	66	53	71	60	77	65,4	Tidak
157	Isnainiyah	Baru	2019	80	71	78	64	69	72,4	Lulus
158	Lum'atul Himmah	Baru	2019	68	75	75	50	85	70,6	Lulus
159	Nadia Hilyat Aulia	Baru	2019	68	85	77	67	75	74,4	Lulus
160	Devatur Rifqiyah	Lama	2019	58	64	71	62	73	65,6	Tidak
161	Khoirun Nisa'	Baru	2019	58	85	78	64	68	70,6	Lulus
162	Farida Nurul Aisyah	Lama	2019	85	80	83	75	70	78,6	Lulus
163	Siti Kholifah	Baru	2019	70	63	70	65	65	66,6	Tidak
164	Shinta Nur Kamila	Baru	2019	88	60	58	61	57	64,8	Tidak
165	Fina Wildania Sholeha	Baru	2020	90	81	90	75	70	81,2	Lulus
166	Sofiatul Hasanah	Baru	2020	77	78	81	72	86	78,8	Lulus
167	Qurrota A'yun	Baru	2020	81	80	86	70	65	76,4	Lulus
168	Siti Kholidiyah	Baru	2020	68	71	67	74	72	70,4	Lulus

Lanjutan Dataset

169	Firdausyiyah	Baru	2020	62	76	66	70	58	66,4	Tidak
170	Diana Safitri	Baru	2020	84	78	70	84	73	77,8	Lulus
171	Diva Safarina Aprilia	Lama	2020	70	68	72	69	77	71,2	Lulus
172	Fira Maghfirotur Rizkiyah	Lama	2020	86	76	72	85	64	76,6	Lulus
173	Riskiyatul Katsiroh	Baru	2020	87	80	84	82	65	79,6	Lulus
174	Murtasiyah	Lama	2020	82	80	78	80	76	79,2	Lulus
175	Silvia Meila Sari	Lama	2020	84	79	64	78	67	74,4	Lulus
176	Hilmi Zainatul M	Lama	2020	75	68	79	80	62	72,8	Lulus
177	Rifatul Afifah	Baru	2020	86	80	77	80	70	78,6	Lulus
178	Maziyatul Husna	Baru	2020	71	85	76	85	80	79,4	Lulus
179	Siti Lailiyah	Baru	2020	85	79	67	76	66	74,6	Lulus
180	Nur Fatimah	Lama	2020	85	79	83	75	67	77,8	Lulus
181	Firda Amalia Fatma N	Lama	2020	73	80	77	85	68	76,6	Lulus
182	Sakinatul Fuada	Lama	2020	90	79	89	91	80	85,8	Lulus
183	Kinanti Diyah Lestari	Lama	2020	74	80	82	80	70	77,2	Lulus
184	Qonitatun Nurul F	Lama	2020	77	75	65	71	76	72,8	Lulus
185	Nurul Rosita Putri	Lama	2020	88	83	80	78	82	82,2	Lulus
186	Husna L Muna	Lama	2020	90	80	95	75	64	80,8	Lulus
187	Isyfina Choirun Nisa'	Baru	2020	71	78	66	85	68	73,6	Lulus
188	Annisa Dinda Ayu Maharani	Lama	2020	65	89	85	85	69	78,6	Lulus
189	Latifatul Fauziyah	Baru	2020	72	87	75	80	68	76,4	Lulus
190	Ida Mawaddah	Baru	2020	80	88	90	75	61	78,8	Lulus
191	Dewi Ayu Ainur Rohmah	Baru	2020	65	90	88	90	66	79,8	Lulus
192	Aisyah Debi Damayanti	Lama	2020	55	90	92	80	60	75,4	Lulus
193	Lisa Wahyuni	Lama	2020	77	63	87	72	65	72,8	Lulus
194	Uswatul Qonita	Lama	2020	70	85	88	90	63	79,2	Lulus
195	Atika Rihlah Nagiyah	Lama	2020	75	90	86	80	85	83,2	Lulus
196	Anisa Uzzakiyah	Baru	2020	71	78	66	80	68	72,6	Lulus
197	Mutmainnah	Baru	2020	70	89	71	68	76	74,8	Lulus
198	Nurin Nida Diana	Lama	2020	73	90	81	90	82	83,2	Lulus
199	Zeda Al-widad	Baru	2020	65	61	75	68	80	69,8	Tidak
200	Rizkiya Dwi Novem Wulandari	Baru	2020	67	73	75	77	69	72,2	Lulus
201	Siti Noor Halimah	Lama	2020	75	88	79	67	83	78,4	Lulus
202	Istifadatul Hasanah	Baru	2020	70	90	88	86	82	83,2	Lulus
203	Diana Oktavia Khosiati	Baru	2020	80	95	79	80	84	83,6	Lulus
204	Siti Komaria	Baru	2020	60	90	77	80	71	75,6	Lulus
205	Faizatun Nafi'ah	Baru	2020	70	85	80	78	88	80,2	Lulus
206	Juwita	Baru	2020	75	82	84	87	90	83,6	Lulus
207	Ayuna Sofiatul Maulani	Baru	2020	85	80	75	85	87	82,4	Lulus
208	Fitriana	Lama	2020	80	80	88	75	80	80,6	Lulus
209	Yayi Wulandari	Baru	2020	65	95	80	75	83	79,6	Lulus
210	Sania	Baru	2020	82	89	79	81	80	82,2	Lulus
211	Mar'atus Sholihah	Baru	2020	75	80	85	79	85	80,8	Lulus
212	Wilda Muyassaroh	Baru	2020	73	58	83	80	70	72,8	Lulus
213	Holifah	Lama	2020	63	76	83	72	74	73,6	Lulus
214	Nuri Afidatul Fitriah	Baru	2020	75	78	83	79	88	80,6	Lulus
215	Noor Aini	Baru	2020	77	83	77	85	87	81,8	Lulus
216	Izzah Nur Nabila	Baru	2020	72	76	80	88	85	80,2	Lulus

Lanjutan Dataset

217	Nahdia Bella	Baru	2020	70	71	71	77	88	75,4	Lulus
218	Siti Nur Inayah	Baru	2020	85	70	86	90	90	84,2	Lulus
219	Marisakh Diana Martabillah	Baru	2020	65	85	77	73	82	76,4	Lulus
220	Rohatin	Baru	2020	81	70	74	78	81	76,8	Lulus
221	Putri Sholehati	Baru	2020	67	71	70	69	65	68,4	Tidak
222	Nurul Fauziah	Baru	2020	75	87	83	90	71	81,2	Lulus
223	Tesa Murni Diah	Baru	2020	68	74	69	80	75	73,2	Lulus
224	Nurani Berliana	Baru	2020	56	77	68	71	75	69,4	Tidak
225	Rofifatul Jannah	Baru	2020	50	64	60	57	65	59,2	Tidak
226	Ariska Anggraini	Baru	2020	75	71	82	85	68	76,2	Lulus
227	Innani Fariyah F	Baru	2020	69	72	70	75	63	69,8	Tidak
228	Nadia Ainun Nufus	Baru	2020	75	78	85	80	70	77,6	Lulus
229	Qomaria	Baru	2020	75	63	73	55	75	68,2	Tidak
230	Latifatul Akmalia	Baru	2020	85	72	81	83	76	79,4	Lulus
231	Fitria Eka Putri Rinjani	Baru	2020	71	70	73	60	75	69,8	Tidak

4.2. Preprocessing Data

Proses ini berfungsi sama seperti reduksi data yaitu untuk mengurangi jumlah volume data dengan pertimbangan hasil akhir yang mendekati pola yang diinginkan dengan cara *Smoothing* untuk menghilangkan data *noisy*, *Attribute Construction* untuk membuat *attribute* baru yang memiliki karakteristik yang sama, *Aggregate Data*, *Normalisasi Data*, *Discretization Data*. Yang terakhir adalah *Integration Data* atau Integrasi Data yaitu proses kombinasi dari beberapa sumber data menjadi satu *database / dataset*. *Preprocessing Data* ini memiliki keunggulan seperti memperlancar proses *data mining*, data agar lebih mudah dimengerti, beban representasi dalam data dapat berkurang, mengurangi waktu *data mining*, dan proses *Machine Learning* dan analisis data dapat dengan mudah dilakukan. Untuk batasan penilaian dalam setiap variabel dan data santri perangkatan dapat dilihat pada tabel dibawah ini

Tabel.4.2. Nilai Tajwid

Tajwid	Keterangan
Nilai > 70	1 (Lulus)
Nilai ≤ 70	0 (Tidak Lulus)

Tabel 4.3. Nilai Baca Tulis Al-qur'an

BTQ	Keterangan
Nilai > 70	1 (Lulus)
Nilai ≤ 70	0 (Tidak Lulus)

Tabel 4.4. Nilai Furdul 'Ainiyah

Furudul 'Ainiyah	Keterangan
Nilai > 70	1 (Lulus)
Nilai ≤ 70	0 (Tidak Lulus)

Tabel 4.5 Nilai Fiqih

Fiqih	Keterangan
Nilai > 70	1 (Lulus)
Nilai ≤ 70	0 (Tidak Lulus)

Tabel 4.6 Nilai Aqidah

Aqidah	Keterangan
Nilai > 70	1 (Lulus)
Nilai ≤ 70	0 (Tidak Lulus)

Tabel 4.7 Nilai Ujian Akhir

Ujian Akhir	Keterangan
Nilai > 70	1 (Lulus)
Nilai ≤ 70	0 (Tidak Lulus)

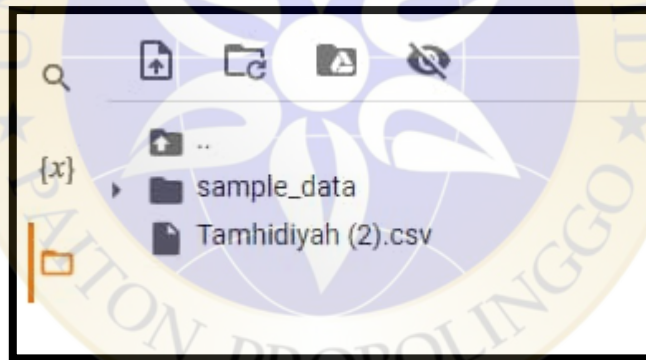
Tabel 4.8 Data Santri Perangkatan

Angkatan	Lama	Baru
2018	49	36
2019	51	27
2020	45	21

4.3. Implementasi *K-Nearest Neighbor*

Setelah tahap-tahap sebelumnya sudah dilakukan maka dilanjut ke dalam tahap selanjutnya yaitu tahap implementasi algoritma *K-Nearest Neighbor* dengan menggunakan Google Colaboratory sebagai media analisisnya. Dengan jumlah data santri yang digunakan sebanyak 231.

- Langkah pertama yang harus kita lakukan adalah mengunggah atau mengupload terlebih dahulu dataset yang akan kita gunakan ke dalam Google Colaboratory.



Gambar 4.1 File tersimpan

Segmen Program 4.1 Import *Library*

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
```

Library yang digunakan disini ada *numpy*, *matplotlib*, *pandas*, dan *seaborn*. *Pandas* sendiri dapat diartikan sebagai *library* yang bisa digunakan untuk menganalisis data pada Python. Sedangkan *matplotlib* adalah *library* yang dapat digunakan untuk visualisasi data. Dan *seaborn* merupakan *library* yang digunakan untuk menghasilkan grafik statistik dari kumpulan data.

Segmen Program 4.2 Input data csv

```
[ ] dataset = pd.read_csv('Tamhadiyah (2).csv')
```

Gambar 4.2 membaca data

Pada segmen 4.2 ini adalah untuk membaca data yang kita upload sebelumnya pada segmen 4.1

```
[ ] dataset.head()
```

	STATUS	TAJWID	BTQ	FA	FIQIH	AQIDAH	Nilai Akhir	Kelas
0	Lama	87	85	80	77	82	82.2	1
1	Baru	63	70	54	64	65	63.2	0
2	Baru	95	74	85	63	91	81.6	1
3	Baru	85	70	81	77	80	78.6	1
4	Baru	71	75	64	53	76	67.8	0

Gambar 4.3 Menampilkan 5 Data Teratas

Selanjutnya menuju proses pembagian data kedalam x dan y, yang dilakukan pada segmen 4.3

Segmen Program 4.3 Membaca Data dalam kolom

```
[ ] x = dataset.iloc[:,[3,6]].values
    y = dataset.iloc[:, -1].values
```

Data x disini yang kita panggil adalah kolom ketiga dan keenam, dimana kolom tersebut berisi nilai FA dan Nilai Ujian Akhir. Pada segmen ini kita menggunakan *iloc* dimana fungsi *iloc* adalah sebagai lokasi yang kita gunakan.

Segmen Program 4.4 membaca isi data x

```
[ ] print(x)
[[80.  82.2]
 [54.  63.2]
 [85.  81.6]
 [81.  78.6]
 [64.  67.8]
 [68.  69.2]
 [85.  82.6]
 [63.  65.8]
 [75.  79.8]
 [85.  78.4]
 [76.  81.2]
 [74.  71.4]
 [91.  79.4]
 [82.  79.8]
 [73.  80.6]
 [81.  74.6]
 [61.  66.4]
 [73.  76.6]
 [66.  56.8]
 [75.  76.8]
 [90.  84.8]
```

Segmen Program 4.4 ini menampilkan isi data x untuk membuktikan data kita sesuai atau tidak dengan pembagian data yang dilakukan pada segmen Program 4.3 Pembagian Data diatas

Dari 231 jumlah data yang dimiliki, data yang digunakan adalah 173 (75%) data training dan banyak data testing 58 (25%).

Segmen Program 4.7 Klasifikasi K-NN

```
from sklearn.neighbors import KNeighborsClassifier
classifier = KNeighborsClassifier(n_neighbors = 9, metric = 'euclidean', p=2)
classifier.fit(x_train, y_train)

KNeighborsClassifier(metric='euclidean', n_neighbors=13)
```

Pada segmen program 4.7 ini adalah penentuan nilai k , dimana nilai k yang ditetapkan dalam penelitian ini adalah $k=9$ sedangkan untuk perhitungannya menggunakan metode *euclidean*, yang mana metode *euclidean* tersebut adalah sebuah metode yang paling sering digunakan untuk menghitung jarak ke tetangga terdekat.

Segmen Program 4.8 Uji Coba

```
[43] from sklearn.metrics import confusion_matrix
      from sklearn.metrics import accuracy_score

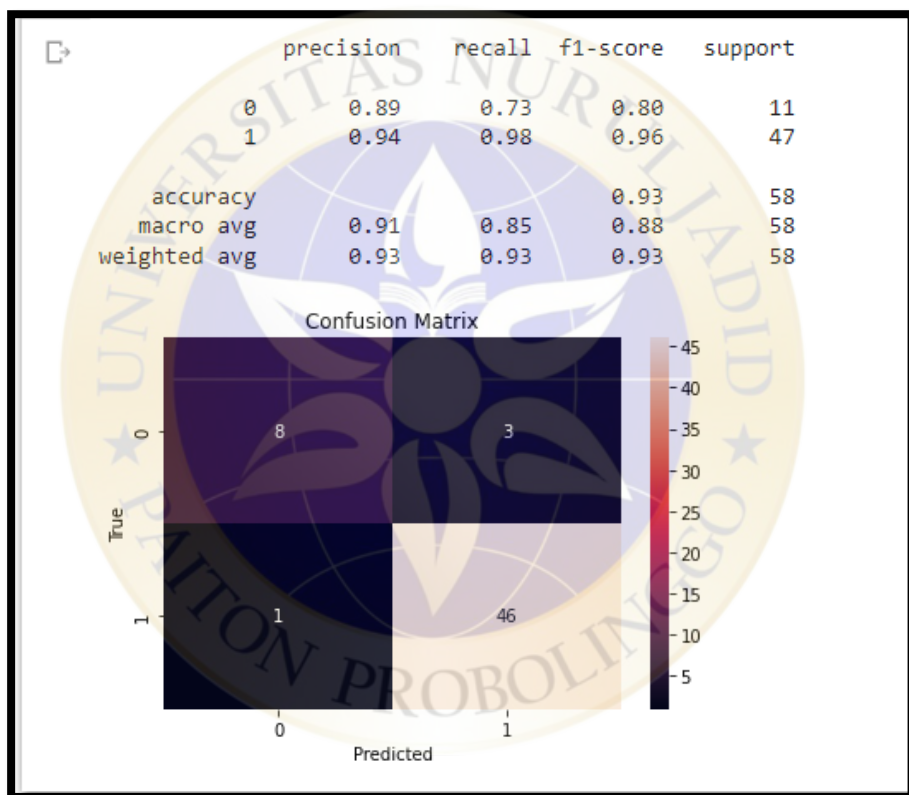
      print(confusion_matrix(y_test, y_pred))
      akurasi=accuracy_score(y_test, y_pred)
      print("Tingkat Akurasi:%d persen"%(akurasi*100))

[[ 8  3]
 [ 1 46]]
Tingkat Akurasi:93 persen
```

Segmen Uji coba ini adalah tahap evaluasi atau validasi menggunakan *confusion matrix* yang nantinya akan digambarkan dengan tabel yang menyatakan jumlah data uji yang benar dan jumlah data uji yang salah diklasifikasikan. Dan untuk hasil yang diperoleh dari evaluasi ini adalah hasil yang menyatakan bahwa santri yang dinyatakan lulus ada 46, dan santri yang dinyatakan tidak lulus ada 8. Dan 3 santri dinyatakan lulus namun diprediksi tidak lulus dan sebanyak 1 santri tidak lulus namun diprediksi lulus. Dengan hasil nilai akurasi yang diperoleh adalah 93%

Segmen Program 4.6 Validasi Hasil

```
#Confusion matrix and classification report
from sklearn import metrics
from sklearn.metrics import classification_report, confusion_matrix
matrix = confusion_matrix(y_test, prediction_classifier)
sns.heatmap(matrix, annot=True, fmt="d")
plt.title('Confusion Matrix')
plt.xlabel('Predicted')
plt.ylabel('True')
print(classification_report(y_test, prediction_classifier))
```

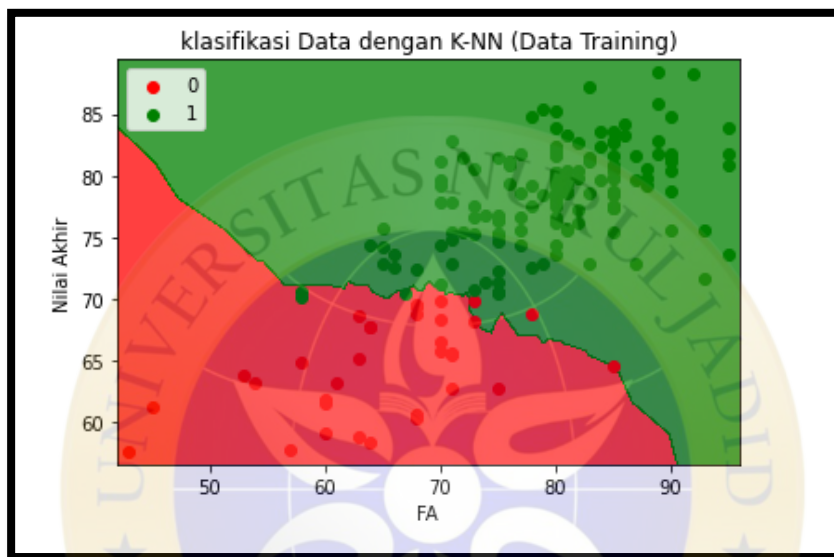


Gambar 4.4 Hasil *confusion_matrix*

confusion_matrix merupakan tabel yang digunakan untuk menggambarkan seberapa baik kinerja model klasifikasi pada sekumpulan data uji. Dan untuk *classification_report* merupakan hasil klasifikasi untuk mengukur kualitas prediksi dari algoritma klasifikasi. Dari hasil validasi data menggunakan *confusion matrix* maka didapat sebuah hasil yang menyatakan bahwa santri yang dinyatakan tidak lulus ada 8. Dan 3 santri yang berstatus lulus namun diprediksi tidak lulus sedangkan 1 santri yang berstatus tidak lulus

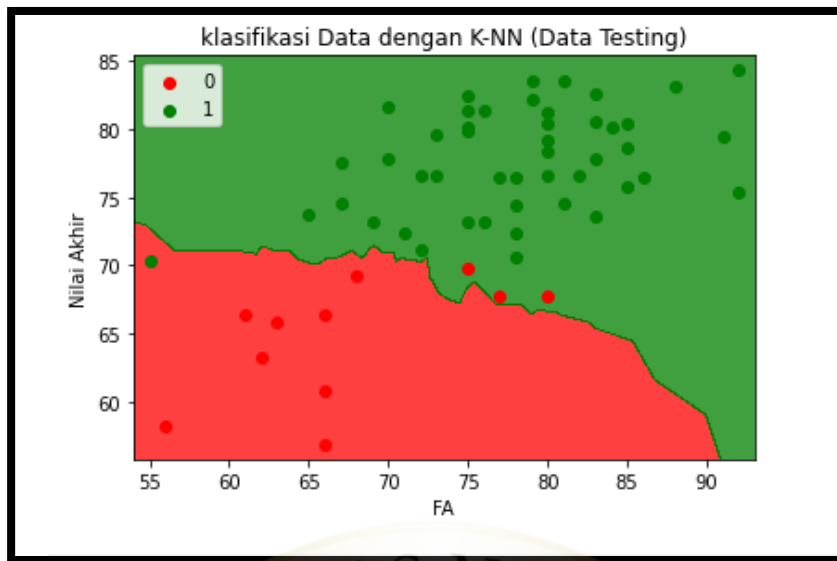
namun diprediksi lulus. Dengan hasil nilai akurasi yang diperoleh adalah 93%. Hasil dari grafik diatas menunjukkan bahwa santri yang berstatus lulus dinyatakan dengan angka 1 dengan *Precision* 94%, *recall* 98% dan *F1-score* sebesar 98% sedangkan santri yang berstatus tidak lulus dinyatakan dengan angka 0 *Precision* 89%, *recall* 73% dan *F1-score* sebesar 80%. Untuk persebaran datanya bisa dilihat pada segmen dibawah ini

Segmen Program 4.7 Visualisasi grafik terhadap data *training* dan data *testing*



Gambar 4.5 Hasil visualisasi data *training*

Dari gambar persebaran data *training* diatas gambar yang berwarna hijau merupakan daerah yang berisi jumlah persebaran data santri yang dinyatakan lulus sedangkan gambar yang berwarna merah merupakan daerah yang berisi jumlah persebaran data santri yang tidak lulus. Untuk nilai Akhir, nilai terbesar adalah 85. Sedangkan untuk FA nilai terbesarnya adalah 90.



Gambar 4.6 Hasil visualisasi data *testing*

Dari gambar 4.6 Hasil visualisasi data *testing* dan data *training* nilai terbesar dari Nilai Akhir adalah 73. Sedangkan untuk nilai FA adalah 90. Berdasarkan gambar diatas ada beberapa santri yang dinyatakan tidak lulus namun diprediksi lulus. Dan ada juga yang lulus namun diprediksi tidak lulus sesuai dengan titik yang ada pada daerah tersebut.

4.4. Hasil Uji Coba

Hasil dari penelitian ini adalah implementasi *data mining* dengan metode *K-Nearest Neighbor* dengan perhitungan jarak dengan menggunakan metode *Euclidean Distance*. Adapun *software* pendukung yang digunakan untuk mengoptimalkan kinerja algoritma *K-Nearest Neighbor* pada penelitian ini adalah *microsoft Excel* dan *Google Colaboratory (Google Colab)*. Dari hasil uji coba yang dilakukan didapatkan hasil yang bagus yaitu 93%. Dataset yang digunakan pada penelitian ini adalah data yang diperoleh dari bagian KA TU Lembaga Tamhadiyah Pondok Pesantren Nurul Jadid Wilayah Al-Hasyimiyah dengan jumlah 231 data dari tahun 2018 -2020. Data tersebut kemudian dibagi menjadi 75% data training dan 25% data testing. Hasil yang didapat dari proses evaluasi dan validasi dengan menggunakan *confusion matrix* yaitu, hasil *Accuracy* sebesar 93%, nilai *preccision* 89%, nilai *reccal* sebesar 73%, dan *f1-score* adalah 80%. Sehingga dapat ditarik kesimpulan bahwasanya algoritma *K-Nearest Neighbor* masuk dalam kategori klasifikasi yang sangat baik.

