

Pengembangan Solusi Perawatan Kesehatan Terhadap *Autism Spectrum Disorder* (ASD) Menggunakan Pendekatan *Data Analysis*

Sitti Rahmah Jabir^{a,1,*}, A.Ulfah Tenripada^{a,2}, Muhammad Arfah Asis^{a,3}, Dewi Widyawati^{a,4}, Amaliah Faradibah^{a,5}

^a Fakultas Ilmu Komputer, Universitas Muslim Indonesia, Jalan Urip Sumoharjo, Makassar, 90231, Indonesia
¹ rahmahjabir@umi.ac.id; ² a.ulfah@umi.ac.id; ³ muh.arfah.asis@umi.ac.id; ⁴ dewiwidyawati@umi.ac.id; ⁵ amaliah.faradibah@umi.ac.id;
*corresponding author

INFORMASI ARTIKEL

Diterima : 14 – 05 – 2022
Direvisi : 23 – 05 – 2022
Diterbitkan : 30 – 05 – 2022

Kata Kunci:
Autism
Data Pre-Processing
Data Transformation
Data Imputation

ABSTRAK

Autisme memiliki masalah dengan interaksi sosial. Mereka tidak dapat mengembangkan hubungan dengan orang lain sesuai dengan tingkat perkembangan mereka. Jumlah anak-anak dengan autisme telah tumbuh terus menerus selama beberapa tahun. Mendiagnosis ASD diperlukan pendekatan yang komprehensif, sistematis, dan terstruktur. Untuk mendiagnosis ASD, peneliti memanfaatkan penambangan data untuk menganalisis data terapi perilaku. Data yang didapatkan tidak sepenuhnya data yang bersih, dimana terdapat beberapa data yang hilang. Untuk menangani data yang hilang, pendekatan *data pre-processing* yang akan digunakan untuk membantu menganalisis dan memperhitungkan nilai yang hilang. Data yang tidak sesuai format akan ditransformasikan terlebih dahulu sebelum divisualisasikan. Sebagian besar kuesioner telah diisi oleh orang tua. Berdasarkan dataset, anak-anak dengan ASD didominasi oleh laki-laki. Dirujuk dari etnis, orang kulit putih-Eropa adalah etnis terbanyak yang terdeteksi memiliki jumlah anak tertinggi dengan ASD. Di dalam etnis, ada berbagai negara. Inggris adalah jumlah terbesar orang yang menderita autisme. Berdasarkan hasil tersebut, bidang kesehatan harus lebih fokus memberikan pengobatan untuk orang kulit putih-Eropa terutama di Inggris. Para peneliti kesehatan harus menghasilkan wawasan yang dapat mengembangkan autisme untuk deteksi dan skrining. Berdasarkan hasil tersebut dapat membantu lebih lanjut agar dapat mengurangi persentase autisme di seluruh dunia. peneliti kesehatan harus menghasilkan wawasan yang dapat mengembangkan autisme untuk deteksi dan skrining.

This is an open access article under the [CC-BY-SA](#) license



I. Pendahuluan

Autisme adalah diagnosis yang mencakup defisit komunikasi sosial, perilaku, minat, atau aktivitas yang terbatas dan berulang yang cukup buruk untuk secara signifikan merusak fungsi [1]. Gangguan spektrum autism, atau biasa disebut *autism spectrum disorder* (ASD), adalah sekelompok kondisi perkembangan saraf. Mereka tidak dapat mengembangkan hubungan dengan orang lain sesuai dengan tingkat perkembangan mereka. Jumlah anak autis terus bertambah selama beberapa tahun [2]. Sebagai gangguan perkembangan saraf yang tidak dapat disembuhkan, ASD harus dideteksi sejak dini. Deteksi dini tidak hanya membantu orang tua untuk mempersiapkan anak-anak mereka juga membantu untuk merancang perangkat Pendidikan. Dimana hal tersebut membantu anak memperoleh Pendidikan yang sesuai [3].

ASD dikategorikan sebagai gangguan perkembangan saraf menurut Diagnostic and Statistical Manual of Mental Disorders. Berdasarkan instrumen angket yang ada, mengacu dari angket yang dilaporkan orang tua yang akan digunakan seluruhnya atau digabungkan dengan skala observasi. *Modified Checklist for Autism in Toddlers* (M-CHAT) / Modifikasi Daftar Periksa Autisme pada Balita adalah instrumen yang paling banyak digunakan dalam pengaturan komunitas dan dievaluasi secara internasional [4]. Pada M-CHAT, terdiri dari item alat skrining ASD laporan orang tua (Ya/Tidak) yang dibuat khusus untuk anak-anak usia 16 hingga 30 bulan. Anak-anak yang terdeteksi autisme pada M-CHAT akan diberikan penilaian diagnostik gratis. Ini diselenggarakan oleh psikolog klinis berlisensi [5].

Jumlah anak-anak yang berada dalam kondisi autisme telah meningkat berulang-ulang dalam beberapa dekade terakhir. Merujuk dari data Centers for Disease Control, saat ini ASD menyerang 1 dari 59 anak di Amerika Serikat [1]. Secara umum, anak ASD menggunakan waktunya untuk bermain pasif. Mereka tidak tertarik untuk terlibat dalam kegiatan rekreasi yang terorganisir seperti olahraga. Keterbatasan bersosialisasi dan berperilaku untuk ASD dapat menghambat kesempatan anak untuk mengikuti kegiatan fisik dan program rekreasi. Selain itu, anak-anak hanya berhenti untuk melakukan aktivitasnya. Salah satu penelitian meneliti 83 anak yang terdiri dari 52 anak laki-laki dan 31 anak perempuan berusia 6-15 tahun yang terdeteksi ASD. Hasil penelitian menunjukkan bahwa hanya 10 (12%) anak ASD yang aktif secara fisik. Sebagian besar dari mereka lebih banyak bermain sendiri daripada melakukan kegiatan bermain social [2].

Saat ini, sebagian besar penelitian berfokus secara khusus pada penggunaan IT yang merupakan pengaruh teknologi informasi untuk melatih anak autisme atau untuk meningkatkan komunikasi dengan mereka. Salah satu penelitian menggunakan pendekatan data mining untuk menganalisis dataset. Salah satu teknik data mining yang akan digunakan dalam penelitian ini adalah klasifikasi. Klasifikasi dapat menjadi proses dua langkah yang terdiri dari data pelatihan dan data pengujian. Pertama, model dibangun dengan memeriksa dan menafsirkan tupel data dalam dataset yang berisi banyak atribut. Kedua, data yang diuji digunakan untuk memeriksa keakuratan model. Jika akurasinya sesuai, maka model dapat digunakan untuk mengklasifikasikan tupel data yang tidak diketahui [6].

Untuk menangani data yang hilang, pendekatan pembelajaran mesin akan digunakan untuk membantu menganalisis dan memasukkan nilai yang hilang. Salah satu teknik yang akan digunakan dalam penelitian ini adalah pohon keputusan. Data yang memiliki jumlah kuantitas tertinggi akan digunakan untuk menghitung data yang hilang. Pohon keputusan menghubungkan nilai dengan nilai sebenarnya. Label kelas telah ditentukan di setiap simpul daun dengan suara terbanyak [7].

Dalam penelitian ini akan digunakan dataset anak autis. Dataset yang diberikan bukanlah data yang jelas. Itu harus dibersihkan sebelum menerapkan teknik data mining untuk mendapatkan solusi yang lebih baik. Tujuan dari penelitian ini adalah untuk mengembangkan hasil perawatan kesehatan untuk membantu anak-anak dengan gangguan spektrum autisme (ASD).

II. Metode

A. Data Imputation

Untuk menangani data yang hilang, pendekatan machine learning memiliki berbagai algoritma untuk membantu analisis data. Dalam penelitian ini akan digunakan metode pohon keputusan dan K-Nearest Neighbor. Pendekatan pohon keputusan adalah label kelas dengan suara terbanyak. Untuk K-Nearest Neighbor (KNN), missing value diisi dengan menghitung rata-rata data lain yang tidak missing.

B. Data Exploration

Eksplorasi data merupakan langkah awal dalam analisis data yang melibatkan perangkat lunak manual dan otomatis yang memvisualisasikan dan mengidentifikasi hubungan antara variabel atau fitur yang berbeda, struktur kumpulan data, keberadaan outlier dalam data, dan distribusi nilai untuk mengungkapkan poin dan pola minat kami memungkinkan analisis untuk mendapatkan wawasan yang lebih baik darinya [8].

C. Data Pre-Processing

Pre-processing merupakan proses awal yang akan mentransformasikan data masukan menjadi data dengan format yang sesuai dan siap untuk diproses. Beberapa contoh hal yang dilakukan dalam *pre-processing* meliputi berbagai proses yang diperlukan antara lain penggabungan, perubahan bentuk, ataupun pentransformasian data sebagai cara untuk membersihkan, mengintegrasikan, mereduksi dan mendiskritisasi.

D. Data Transformation

Transformasi data ialah merubah skala data kedalam bentuk lain sehingga data memiliki distribusi yang diharapkan. Langkah ini diperlukan untuk proses data mining. Beberapa teknik untuk transformasi data diantaranya ialah *normalization*, pemilihan atribut, dan *discretization* [9].

III. Hasil dan Pembahasan

A. Data Exploration

Sebelum data yang akan digunakan di analisis, data akan di eksplor terlebih dahulu. Dimana akan melihat kondisi data yang dimiliki.

Alphabetic List of Variables and Attributes					
#	Variable	Type	Len	Format	Informat
10	A10_Score	Num	8	BEST12.	BEST32.
1	A1_Score	Num	8	BEST12.	BEST32.
2	A2_Score	Num	8	BEST12.	BEST32.
3	A3_Score	Num	8	BEST12.	BEST32.
4	A4_Score	Num	8	BEST12.	BEST32.
5	A5_Score	Num	8	BEST12.	BEST32.
6	A6_Score	Num	8	BEST12.	BEST32.
7	A7_Score	Num	8	BEST12.	BEST32.
8	A8_Score	Num	8	BEST12.	BEST32.
9	A9_Score	Num	8	BEST12.	BEST32.
21	Class_ASD	Char	3	\$3.	\$3.
11	age	Num	8	BEST12.	BEST32.
19	age_desc	Char	12	\$12.	\$12.
15	austim	Char	3	\$3.	\$3.
16	contry_of_res	Char	22	\$22.	\$22.
13	ethnicity	Char	17	\$17.	\$17.
12	gender	Char	1	\$1.	\$1.
14	jundice	Char	3	\$3.	\$3.
20	relation	Char	6	\$6.	\$6.
18	result	Num	8	BEST12.	BEST32.
17	used_app_before	Char	3	\$3.	\$3.

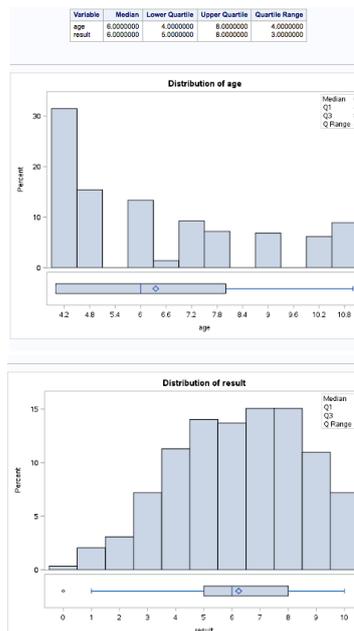
Gambar 1. Dataset

Dari data tersebut ditunjukkan variabel-variabel beserta jenis, panjang, format dan informasinya. Data terdiri dari 21 variabel yang terdiri dari 12 data numerik dan 9 data nominal.

B. Data Pre-processing

1) Outlier

Dimana terdapat outlier yang ditunjukkan pada gambar 2 di bawah ini:



Gambar 2. Grafik Persebaran Data

2) Missing Value

a. Categorical value

Berdasarkan Gambar 1, terdapat adanya nilai yang hilang pada variabel umur. Ada empat baris yang sudah kosong umurnya. Namun, nilai numerik lainnya tidak memiliki nilai yang hilang. Dalam kueri, itu hanya berjalan untuk nilai numerik bukan nilai nominal.

The MEANS Procedure

Variable	N Miss
A1_Score	0
A2_Score	0
A3_Score	0
A4_Score	0
A5_Score	0
A6_Score	0
A7_Score	0
A8_Score	0
A9_Score	0
A10_Score	0
age	4
result	0

Gambar 3. Variabel yang memiliki data yang hilang

Untuk menangani missing value, salah satu metode yang digunakan yaitu menerapkan mean pada data yang hilang. Untuk mengisi nilai yang hilang, diperlukan sarana untuk mengimputasi data. Rerata dari semua nilai umur adalah 6.3541667. Untuk pengisian data miss hanya membutuhkan waktu 6.35.

The MEANS Procedure

Analysis Variable : age				
N	Mean	Std Dev	Minimum	Maximum
288	6.3541667	2.3654563	4.0000000	11.0000000

Gambar 4. Nilai rata-rata variable umur

Data yang hilang akan diisi menggunakan nilai 6.35. Setelah imputasi, *query* tersebut dijalankan kembali untuk mencari missing value yang ada masih ada atau tidak.

The MEANS Procedure

Variable	N Miss
A1_Score	0
A2_Score	0
A3_Score	0
A4_Score	0
A5_Score	0
A6_Score	0
A7_Score	0
A8_Score	0
A9_Score	0
A10_Score	0
age	0
result	0

Gambar 5. Variabel yang memiliki data yang hilang

Selain itu, juga dapat dilihat untuk kategori usia telah bertambah dengan nilai baru.

The FREQ Procedure

age	Frequency	Percent	Cumulative Frequency	Cumulative Percent
4	92	31.51	92	31.51
5	45	15.41	137	46.92
6	39	13.36	176	60.27
6.35	4	1.37	180	61.64
7	27	9.25	207	70.89
8	21	7.19	228	78.08
9	20	6.85	248	84.93
10	18	6.16	266	91.10
11	26	8.90	292	100.00

Gambar 6. Kategori variable umur

b. Nominal value

Setelah semua nilai numerik telah diperhitungkan, langkah selanjutnya adalah untuk nilai nominal. Kategori dalam variable akan ditunjukkan untuk melihat variable yang memiliki missing value pada data dengan kategori nominal.

The FREQ Procedure

Class_ASD	Frequency	Percent	Cumulative Frequency	Cumulative Percent
NO	151	51.71	151	51.71
YES	141	48.29	292	100.00

age_desc	Frequency	Percent	Cumulative Frequency	Cumulative Percent
4-11 years	292	100.00	292	100.00

ethnicity	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Others	14	4.79	14	4.79
Middle Eastern	27	9.25	41	14.04
?	43	14.73	84	28.77
White-European	108	36.99	192	65.75
Black	14	4.79	206	70.55
South Asian	21	7.19	227	77.74
Asian	46	15.75	273	93.49
Pasifika	2	0.68	275	94.18
Hispanic	7	2.40	282	96.58
Turkish	2	0.68	284	97.26
Latino	8	2.74	292	100.00

gender	Frequency	Percent	Cumulative Frequency	Cumulative Percent
m	208	71.23	208	71.23
f	84	28.77	292	100.00

jundice	Frequency	Percent	Cumulative Frequency	Cumulative Percent
no	212	72.60	212	72.60
yes	80	27.40	292	100.00

relation	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Parent	214	73.29	214	73.29
?	43	14.73	257	88.01
Self	4	1.37	261	89.38
Relati	17	5.82	278	95.21
Health	13	4.45	291	99.66
self	1	0.34	292	100.00

used_app_before	Frequency	Percent	Cumulative Frequency	Cumulative Percent
no	281	96.23	281	96.23
yes	11	3.77	292	100.00

Gambar 7. Kategori variable nominal

Dari hasil tersebut terlihat nilai '?' yang terdapat pada dataset. Oleh karena itu, terdapat *missing value* yang diberi tanda '?'. Tanda tersebut terdapat pada variabel etnis dan relasi. Untuk suku dan relasi, mereka memiliki kuantitas yang sama untuk simbol tanda tanya. Kedua variabel tersebut terdiri dari 43 nilai yang tidak diketahui.

Untuk pengisian nilai '?' akan dirujuk dari kolom *contry_of_res*. Data tersebut berisi nama negara yang dapat dijadikan acuan. Nilai yang hilang dalam etnis akan merujuk pada etnis lain yang memiliki negara yang sama. Namun, tidak semua negara memiliki referensi untuk etnis. Untuk lainnya yang tidak memiliki acuan, dapat diisi dengan menggunakan 'Lainnya' yang sudah ada dalam kategori variable.

Table 1. Refrensi Negara berdasarkan *Ethnicity*

<i>Ethnicity</i>	<i>Country</i>
<i>Middle Eastern</i>	Jordan, Syria, Egypt, United Arab Emirates, Iraq, Saudi Arabia, Pakistan, Malaysia, Ausria
<i>South Asian</i>	Saudi Arabia
<i>Asian</i>	Pakistan, Malaysia
<i>White-European</i>	Austria
<i>Others</i>	China, Russia, Latvia, Kuwait, Qatar, Lebanon

Setelah diisi, data akan dicek kembali apakah masih ada *missing value* atau tidak. Dari hasil di bawah ini, terlihat tidak ada lagi nilai '?' untuk etnisitas. Nilai nominal selanjutnya untuk imputasi adalah relasi. Itu mendapat 43 nilai yang tidak diketahui di dalam data. Tidak hanya memiliki nilai yang hilang tetapi juga terdiri dari data yang bising yaitu 'diri' dan 'Diri'. Untuk nilai yang tidak diketahui akan diisi dengan mengacu pada jumlah kategori data tertinggi yaitu 'Parent'. Untuk data *noise*, data akan digabungkan dimana 'self' akan diubah menjadi 'Self' sebagai jumlah tertinggi yang akan dirujuk.

The FREQ Procedure

relation	Frequency	Percent	Cumulative Frequency	Cumulative Percent
?	43	14.73	43	14.73
Health	13	4.45	56	19.18
Parent	214	73.29	270	92.47
Relati	17	5.82	287	98.29
Self	4	1.37	291	99.66
self	1	0.34	292	100.00

Gambar 8. Kategori variable relasi

C. Data Transformation

Untuk transformasi, data kuantitatif akan diganti dengan kualitatif dalam tugas ini. Hal ini dikarenakan beberapa algoritma data mining/machine learning dirancang untuk bekerja dengan data kualitatif atau kuantitatif, campuran hanya dapat didukung oleh beberapa algoritma.

1) Score Data

Semua skor memiliki nilai yang sama. Ini menunjukkan 0 dan 1. Untuk mengubah data, 0 diubah menjadi "Tidak" dan 1 diganti sebagai "Ya".

2) Age Data

Data umur diubah menjadi kelompok umur. Umur terdiri dari 9 atribut yaitu 4, 5, 6, 6.35, 7, 8, 9, 10, 11.

The FREQ Procedure

age	Frequency	Percent	Cumulative Frequency	Cumulative Percent
4	92	31.51	92	31.51
5	45	15.41	137	46.92
6	39	13.36	176	60.27
6.35	4	1.37	180	61.64
7	27	9.25	207	70.89
8	21	7.19	228	78.08
9	20	6.85	248	84.93
10	18	6.16	266	91.10
11	26	8.90	292	100.00

Gambar 9. Kategori variable umur

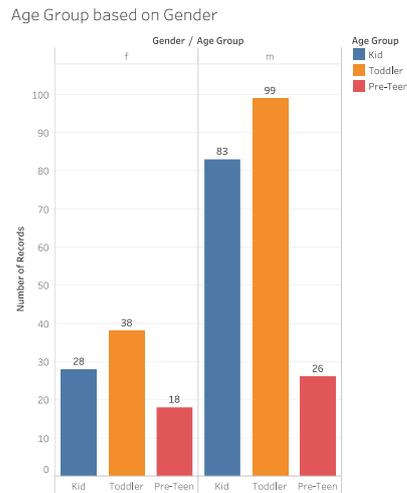
Dibagi menjadi tiga kelompok. Kelompok tersebut adalah Toddler (dibawah 5 tahun), Kid (dari 5 hingga di bawah 8 tahun), dan Pra-Remaja (8 tahun ke atas). Kuantitas masing-masing atribut tidak terlalu berbeda.

age_group	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Kid	115	39.38	115	39.38
Pre-Teen	85	29.11	200	68.49
Toddler	92	31.51	292	100.00

Gambar 10. Kategori baru variable umur

D. Data Visualization

1) Hipotesis I

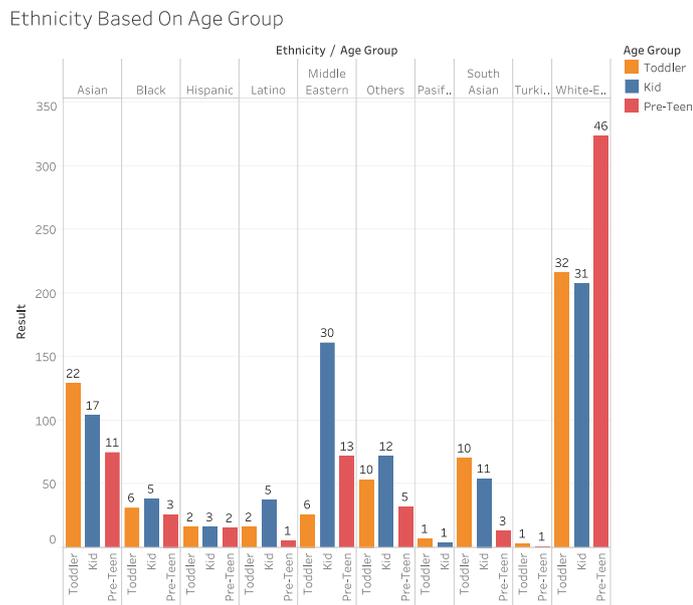


Gambar 11. Persebaran umur (*age*) berdasarkan jenis kelamin (*gender*)

Berdasarkan jenis kelamin, anak laki-laki mendapat persentase kelompok usia tertinggi dibandingkan dengan anak perempuan. Gambar untuk kedua jenis kelamin sama sekali berbeda dengan perempuan dengan ASD kira-kira hanya setengah persentase dari laki-laki. Dirujuk dari penelitian lain, pihaknya menyelidiki 83 anak yang terdiri dari 52 anak laki-laki dan 31 anak perempuan berusia 6-15 tahun yang terdeteksi ASD. Objek lebih dinominasikan oleh laki-laki.

2) Hipotesis II

Eropa adalah etnis yang memiliki jumlah Anak Autisme tertinggi.



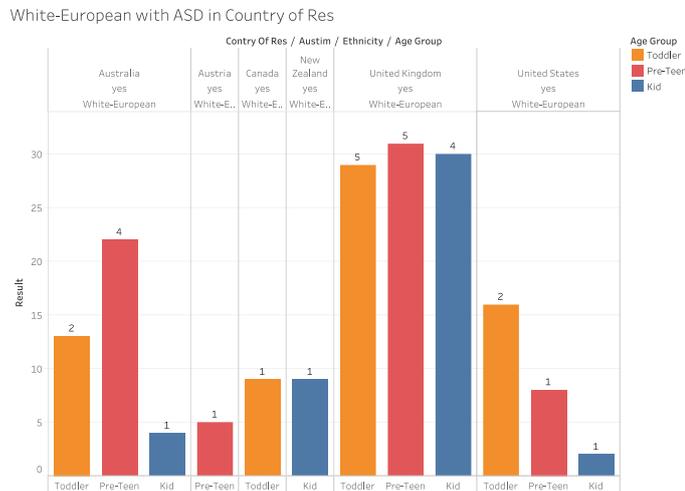
Gambar 12. Persebaran etnis (*ethnicity*) berdasarkan grup umur (*age group*)

Dibandingkan dengan etnis lain, grafik menunjukkan pada etnis kulit putih Eropa, mereka mendapat persentase autisme anak yang tinggi. Grafik yang paling didominasi adalah untuk Pra-Remaja yaitu anak-anak berusia sekitar 8 hingga 11 tahun mendapatkan persentase yang banyak

dibandingkan kelompok usia lainnya. Sebanyak 46 anak tercatat dalam persentase anak autisme tertinggi dibandingkan etnis dalam dataset.

3) Hipotesis III

United Kingdom mendominasi persentase anak autisme untuk orang kulit putih-Eropa.

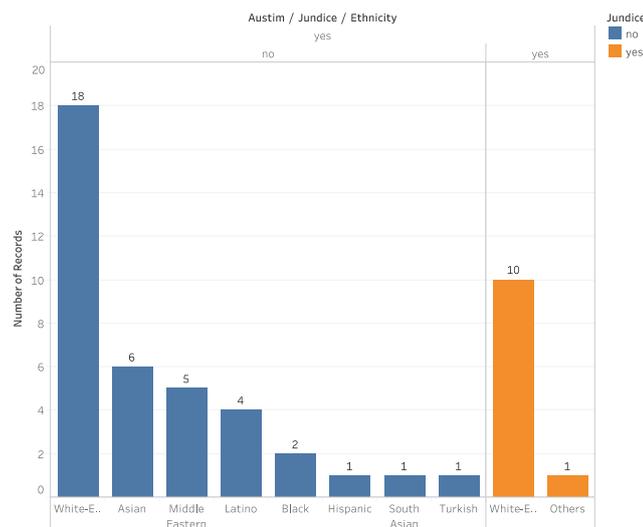


Gambar 13. Persebaran penderita ASD eropa kulit putih berdasarkan negara

Berdasarkan hipotesis sebelumnya, sebagian besar etnis yang memiliki kulit putih-Eropa tertinggi di setiap kelompok umur. Dalam etnis itu, terdiri dari banyak negara. *Bar graph* menunjukkan bahwa Inggris (United Kingdom) adalah negara yang paling banyak terdeteksi dengan ASD.

4) Hipotesis IV

Penyakit kuning tidak sepenuhnya dipengaruhi oleh autisme. Dimana dapat dilihat pada gambar di bawah ini.

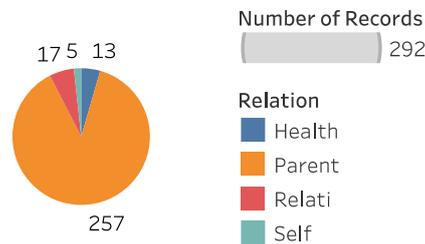


Gambar 14. Persebaran penderita ASD sekaligus penyakit kuning berdasarkan etnis

Hanya ada 2 etnis anak autisme yang mengalami penyakit kuning. Orang kulit putih-Eropa memperoleh jumlah anak yang tinggi dengan ASD bersama dengan penyakit kuning. Namun, hanya setengah dari persentase autisme yang juga memiliki penyakit lain.

5) Hypothesis V

Sebagian besar kuesioner yang disebar, diisi oleh orang tua.

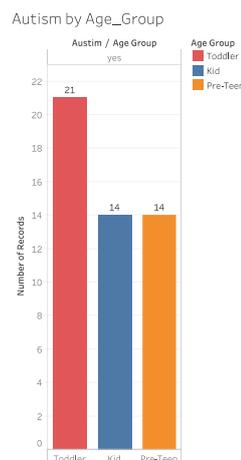


Gambar 15. Persebaran responden kuisisioner

M-CHAT adalah alat skrining ASD laporan orang tua (Ya/Tidak) yang dirancang khusus untuk anak-anak berusia 16 hingga 30 tahun. M-CHAT (Modified Checklist for Autism in Toddlers) adalah instrumen yang paling banyak digunakan dalam pengaturan komunitas dan dievaluasi secara internasional. Oleh karena itu, alat skrining M-CHAT dapat diimplementasikan dalam penelitian di masa depan.

6) Hipotesis 6

Sebagian besar jumlah rekor didominasi oleh Toddler Group.



Gambar 16. Persebaran penderita ASD berdasarkan kelompok umur (*Age Group*)

Anak-anak di bawah 5 tahun terdeteksi autisme lebih dari kelompok lain. Berdasarkan dataset yang berisi data 292 anak, hanya ada 49 dengan autisme. 21 dari 49 data didominasi oleh Balita.

Oleh karena itu, analisis akan dilakukan dengan menggunakan M-CHAT yang alatnya adalah daftar periksa yang dimodifikasi khusus untuk autisme pada balita.

IV. Kesimpulan dan saran

Etnis yang memiliki persentase tertinggi adalah Kulit Putih-Eropa. Sebagian besar orang Kulit Putih-Eropa menderita ASD dibandingkan dengan etnis lain. Di dalam etnis, ada banyak negara yang terlibat. Negara yang menghasilkan jumlah anak dengan autisme yang tinggi adalah Inggris. Untuk jenis kelamin, didominasi oleh anak laki-laki terutama untuk com p ASD adalah Oleh karena itu, pengembangan penelitian kesehatan harus lebih fokus pada anak laki-laki dengan ASD. Para peneliti perlu menemukan pengobatan yang baik untuk anak laki-laki dengan ASD. Meskipun anak laki-laki lebih fokus, itu bukan cara untuk menghindari perawatan wanita. Ini juga membutuhkan pendekatan serta anak laki-laki.

Di sisi lain, kelompok Balita juga harus lebih dieksplorasi. Hal ini disebabkan persentase yang lebih tinggi dari kelompok lain. Mendiagnosis akan lebih mudah dengan beberapa alat untuk menganalisis autisme. Ini

ditawarkan dalam banyak alat. Salah satu alat yang menggunakannya untuk mendiagnosis Balita adalah M-CHAT. Ini dapat digunakan untuk pengaturan komunitas dan dievaluasi secara internasional. Selain itu, beberapa kuesioner diisi oleh orang tua. Oleh karena itu, data tersebut cocok untuk diimplementasikan di M-CHAT sebagai alat skrining.

Daftar Pustaka

- [1] M. Defilippis, "Depression in children and adolescents with autism spectrum disorder," *Children*, vol. 5, no. 9, 2018, doi: 10.3390/children5090112.
- [2] A. H. Memari *et al.*, "Children with Autism Spectrum Disorder and Patterns of Participation in Daily Physical and Play Activities," *Neurol. Res. Int.*, vol. 2015, no. July, 2015, doi: 10.1155/2015/531906.
- [3] B. Ida Seraphim, L. S. Rao, and S. Joshi, "Survey on early detection of autism using data mining techniques," *Int. J. Eng. Technol.*, vol. 7, no. 2, pp. 79–80, 2018, doi: 10.14419/ijet.v7i2.24.12003.
- [4] N. Haglund, S. O. Dahlgren, P. Gustafsson, and ..., "Early screening for autism in child health services," ... *J. Autism ...*, vol. 3, no. 1, pp. 01–07, 2017, [Online]. Available: [http://portal.research.lu.se/portal/en/publications/early-screening-for-autism-in-child-health-services\(ed3e8847-66a5-4bfc-975b-f1644bf97549\)/publications.html](http://portal.research.lu.se/portal/en/publications/early-screening-for-autism-in-child-health-services(ed3e8847-66a5-4bfc-975b-f1644bf97549)/publications.html).
- [5] M. L. McPheeters *et al.*, "Screening for Autism Spectrum Disorder in Young Children: A Systematic Evidence Review for the U.S. Preventive Services Task Force," *AHRQ Publ. No. 13-05185-EF-1*, no. 121, p. 202, 2015, [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/26985520%5Chttp://www.uspreventiveservicestaskforce.org/Page/Document/draft-evidence-review106/autism-spectrum-disorder-in-young-children-screening>.
- [6] M. S. Mythili and A. R. M. Shanavas, "A Study on Autism Spectrum Disorders using Classification Techniques," *Int. J. Soft Comput. Eng.*, no. 5, pp. 2231–2307, 2014, [Online]. Available: <http://www.ijscce.org/wp-content/uploads/papers/v4i5/E2433114514.pdf>.
- [7] Z. A. Nadzurah, I. Amelia Ritahani, and A. Nurul, "Performance Analysis of Machine Learning Algorithms for Missing Value Imputation," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 6, 2018.
- [8] R. Magdum, "What is Data Exploration? and its Importance in Data Analytics," *Int. Res. J. Eng. Technol.*, vol. 09, no. 01, p. 01, 2022.
- [9] Studocu, "Data Transformasi-Big Data," 2014. .