

Model Bayesian Network untuk Analisis Risiko dan Prediksi Kasus Pelecehan Seksual di Perguruan Tinggi Palangka Raya

Patricia Elena Putri¹⁾, Rony Teguh²⁾, Ariesta Lestari³⁾

¹⁾²⁾³⁾Teknik Informatika, Fakultas Teknik, Universitas Palangka Raya Kampus Tunjung Nyaho
Jalan Yos Sudarso, Palangka Raya, Kalimantan Tengah, Indonesia

¹⁾patriciaaelena@mhs.eng.upr.ac.id

²⁾ronyteguh@it.upr.ac.id

³⁾ariesta@it.upr.ac.id

Abstrak

Pelecehan seksual di lingkungan pendidikan tinggi merupakan bentuk kekerasan berbasis gender yang kompleks dan sering kali tidak dilaporkan. Penelitian ini bertujuan untuk menganalisis faktor-faktor risiko serta memprediksi kemungkinan terjadinya pelecehan seksual di perguruan tinggi Kota Palangka Raya menggunakan pendekatan *Bayesian Network*. Data diperoleh dari 304 responden melalui kuesioner yang disusun berdasarkan indikator Permendikbudristek No. 30 Tahun 2021. Proses pra-pemrosesan data mencakup pembersihan data, *balancing* kelas target, *feature engineering*, dan pembentukan *intermediate node* untuk menyederhanakan struktur jaringan. Tiga model diuji dengan metode *Expectation-Maximization* untuk pembelajaran parameter dan dibandingkan menggunakan evaluasi akurasi, *confusion matrix*, dan kurva ROC-AUC. Hasil menunjukkan bahwa model dengan *balancing* data, *intermediate node*, dan *stratified split* (Model C) memberikan performa terbaik, dengan recall 100% dalam mendeteksi korban dan AUC sebesar 0.95. Temuan ini menunjukkan bahwa model dapat digunakan sebagai alat bantu dalam pemetaan risiko pelecehan seksual dan pengambilan kebijakan preventif di lingkungan perguruan tinggi.

Kata kunci: Pelecehan Seksual, *Bayesian Network*, Perguruan Tinggi, Faktor Risiko, *Expectation-Maximization*

Abstract

Sexual harassment in higher education institutions is a complex form of gender-based violence and often goes unreported. This study aims to analyze risk factors and predict the likelihood of sexual harassment occurring in universities in Palangka Raya City using a Bayesian Network approach. Data were collected from 304 respondents through a questionnaire designed based on the indicators of Permendikbudristek No. 30 of 2021. The data preprocessing included data cleaning, target class balancing, feature engineering, and the creation of intermediate nodes to simplify the network structure. Three models were tested using the Expectation-Maximization method for parameter learning and compared using accuracy evaluation, confusion matrix, and ROC-AUC curve. The results show that the model with data balancing, intermediate nodes, and stratified split (Model C) provided the best performance, achieving 100% recall in detecting victims and an AUC score of 0.95. These findings indicate that the model can be used as a decision-support tool for mapping the risk of sexual harassment and formulating preventive policies in higher education environments.

Keywords: Sexual Harassment, Bayesian Network, Higher Education, Risk Factors, Expectation-Maximization.

1. PENDAHULUAN

Pelecehan seksual di lingkungan perguruan tinggi merupakan bentuk kekerasan berbasis gender yang kompleks dan sering kali tidak dilaporkan. Fenomena ini berdampak luas bagi korban, baik secara psikologis, sosial, maupun akademik. Kasus-kasus yang terjadi di perguruan tinggi dapat menyebabkan trauma, penurunan produktivitas belajar, hingga putus studi. Meskipun perempuan sering menjadi korban utama, laki-laki juga bisa mengalami hal serupa, dan pelaku dapat berasal dari berbagai kalangan, termasuk dosen, mahasiswa, maupun tenaga kependidikan.

Di Kota Palangka Raya, kekerasan seksual di institusi pendidikan masih menjadi permasalahan serius. Berdasarkan data Komnas Perempuan dan Sistem Informasi Online Perlindungan Perempuan dan Anak (SIMFONI PPA), tercatat 108 kasus kekerasan seksual di Kota Palangka Raya dalam kurun waktu 2019 hingga 2023. Dua di antaranya merupakan kasus yang melibatkan oknum dosen Universitas Palangka Raya, yang turut mendorong pembentukan Satuan Tugas Pencegahan dan Penanganan Kekerasan Seksual (Satgas PPKS) di institusi tersebut.

Sayangnya, analisis berbasis data untuk memahami pola pelecehan seksual di lingkungan kampus masih terbatas. Hal ini menyulitkan pihak kampus dalam merancang kebijakan pencegahan yang tepat dan berbasis konteks. Untuk itu, dibutuhkan pendekatan analitis yang mampu memetakan keterkaitan antar faktor risiko pelecehan seksual secara sistematis dan kuantitatif.

Beberapa penelitian terdahulu telah membuktikan efektivitas *Bayesian Network* sebagai alat analisis untuk memetakan pola keterkaitan antar variabel yang kompleks dan penuh ketidakpastian. Misalnya, Penelitian yang menggunakan *Bayesian Network* untuk mendukung penilaian terhadap kasus pelecehan seksual anak di Finlandia, menunjukkan bahwa model ini mampu menggabungkan berbagai indikator sosial dan psikologis untuk menghasilkan prediksi yang akurat [1]. Penelitian lain menggunakan pendekatan *Hierarchical Bayesian Model* untuk mengestimasi jumlah kasus kekerasan seksual yang tidak dilaporkan di kampus-kampus di Amerika Serikat. Mereka menemukan bahwa banyak institusi pendidikan cenderung mengalami underreporting, terutama di kampus kecil atau konservatif [2].

Berbagai studi tersebut memperlihatkan bahwa pendekatan Bayesian dapat mengungkap pola tersembunyi dalam data sosial dan membantu dalam pengambilan keputusan berbasis bukti. Hal ini sejalan dengan pendapat [3], bahwa *Bayesian Network* mampu menggambarkan hubungan sebab-akibat antar variabel secara eksplisit melalui struktur *Directed Acyclic Graph* (DAG), serta mendukung penalaran probabilistik dalam sistem cerdas.

Penelitian ini bertujuan membangun model *Bayesian Network* untuk menganalisis faktor-faktor risiko dan memprediksi kemungkinan terjadinya pelecehan seksual di lingkungan perguruan tinggi Kota Palangka Raya. Model dikembangkan berdasarkan data dari 304 responden sivitas akademika yang dikumpulkan melalui kuesioner daring yang disusun berdasarkan indikator Permendikbudristek No. 30 Tahun 2021. Proses pra-pemrosesan data meliputi pembersihan data, *balancing* kelas target, *feature engineering*, serta pembentukan *intermediate node* untuk menyederhanakan struktur jaringan. Proses pemodelan dilakukan menggunakan perangkat lunak *GeNIe Modeler* dan metode *Expectation-Maximization* dalam pembelajaran parameter.

Evaluasi model dilakukan melalui metrik akurasi, *recall*, *precision*, *Area Under Curve* (AUC), dan confusion matrix untuk membandingkan performa dari beberapa struktur model. Hasil dari penelitian ini diharapkan dapat menjadi dasar dalam perumusan strategi pencegahan kekerasan seksual yang lebih terarah dan berbasis data di lingkungan perguruan tinggi.

2. TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Penelitian dari Universitas Abo Akademi, Finlandia, mengembangkan alat bantu berbasis *Bayesian Network* untuk mendukung penilaian kasus pelecehan seksual anak. Model ini menggunakan data dari 10.665 laporan dan mempertimbangkan variabel seperti karakteristik individu, gejala yang dilaporkan, serta latar belakang sosial dan lingkungan. Hasilnya

menunjukkan bahwa model mampu memprediksi risiko pelecehan secara akurat dan efektif digunakan sebagai *decision-support tool* dalam proses investigasi [1]. Penelitian lainnya dari *Columbia University* mengembangkan *Hierarchical Bayesian Model* untuk mengestimasi jumlah kasus kekerasan seksual di perguruan tinggi yang tidak dilaporkan secara resmi. Dengan menganalisis data dari 1.973 institusi pendidikan tinggi di Amerika Serikat, mereka menemukan bahwa tingkat pelaporan aktual jauh lebih rendah dari angka resmi, terutama di institusi kecil atau konservatif. Pendekatan ini menunjukkan kemampuan metode *Bayesian* dalam mengungkap pola tersembunyi dalam data sosial dan mendukung perumusan kebijakan yang lebih adil dan berbasis data [2]. Penelitian dari *Harvard University* menggunakan *Bayesian Network* untuk menganalisis hubungan antar gejala PTSD pada orang dewasa yang mengalami pelecehan seksual di masa kanak-kanak. Dengan memanfaatkan pendekatan graphical LASSO dan directed acyclic graph (DAG), mereka menemukan bahwa gejala seperti reaktivitas fisiologis terhadap trauma, mimpi buruk, dan kehilangan minat merupakan pusat dari jaringan gejala. Studi ini menegaskan bahwa pendekatan jaringan dapat memahami PTSD sebagai sistem yang saling berinteraksi dan membantu merancang intervensi yang lebih efektif [4].

2.2 Pelecehan Seksual di Lingkungan Perguruan Tinggi

Pelecehan seksual merupakan bentuk kekerasan berbasis gender yang melibatkan perilaku atau tindakan seksual yang tidak diinginkan oleh korban. Di lingkungan perguruan tinggi, pelecehan seksual dapat terjadi antara dosen dan mahasiswa, antarmahasiswa, maupun antara tenaga kependidikan dan mahasiswa. Pelecehan seksual tidak selalu dalam bentuk fisik, namun juga bisa berbentuk verbal dan simbolik yang merendahkan martabat korban [5]. Lebih lanjut, pelecehan ini sering kali bersifat sistemik, memanfaatkan relasi kuasa antara pelaku dan korban [6].

Dalam konteks pendidikan tinggi, survei menunjukkan bahwa banyak kasus pelecehan seksual tidak dilaporkan karena rasa takut akan stigma, sanksi akademik, atau tidak tersedianya mekanisme pelaporan yang aman [7]. Kegagalan pada institusi pendidikan dalam menangani kasus pelecehan dapat memperburuk trauma korban dan menurunkan kualitas lingkungan akademik secara keseluruhan [8].

2.3 Faktor Resiko dan Teori Sosial

Terdapat dua faktor utama penyebab pelecehan seksual, yaitu faktor biologis dan sosial budaya. Faktor sosial budaya lebih dominan dalam menjelaskan kecenderungan terjadinya pelecehan, terutama dalam masyarakat yang masih menerapkan sistem patriarki [9]. *Strain Theory* menjelaskan bahwa perilaku menyimpang, termasuk pelecehan seksual, dapat terjadi akibat tekanan sosial yang tidak dapat dipenuhi individu melalui cara yang sah [10].

Di lingkungan perguruan tinggi, tekanan akademik, relasi kuasa antara dosen dan mahasiswa, serta minimnya kontrol sosial dapat menciptakan kondisi yang memicu tindakan menyimpang. Oleh karena itu, pendekatan struktural sangat penting dalam memetakan risiko dan merancang kebijakan pencegahan yang komprehensif.

2.4 Bayesian Network

Bayesian Network merupakan metode pemodelan statistik berbasis graf yang menggambarkan hubungan *probabilistik* antar variabel dalam bentuk *Directed Acyclic Graph* (DAG). Setiap simpul (*node*) dalam jaringan merepresentasikan variabel acak, sementara sisi berarah menunjukkan hubungan sebab-akibat [11]. Menurut Korb dan Nicholson (2010), *Bayesian Network* sangat efektif dalam menyederhanakan sistem kompleks dan memungkinkan inferensi probabilistik berbasis data yang tidak lengkap atau tidak pasti [12].

Komponen utama *Bayesian Network* meliputi struktur jaringan dan tabel probabilitas bersyarat (*Conditional Probability Table/CPT*). Node yang tidak memiliki induk disebut *prior nodes* dan memiliki distribusi probabilitas awal. Sementara itu, untuk menyederhanakan struktur jaringan,

dapat digunakan *intermediate nodes* yang berfungsi sebagai simpul penghubung antar variabel yang kompleks [13].

Kelebihan utama dari *Bayesian Network* adalah kemampuannya melakukan inferensi untuk memprediksi kemungkinan suatu kejadian berdasarkan kondisi atau bukti yang diamati, serta fleksibilitasnya dalam mengakomodasi perubahan struktur data.

3. METODE PENELITIAN

3.1 Jenis dan Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif deskriptif untuk menganalisis pola keterkaitan antar faktor risiko dalam kasus pelecehan seksual di lingkungan perguruan tinggi. Pendekatan ini dipilih karena sesuai untuk mengukur hubungan antar variabel secara sistematis melalui data numerik, serta memungkinkan penerapan metode statistik dan pemodelan probabilistik. Pendekatan kuantitatif tepat digunakan dalam penelitian yang bertujuan menguji hubungan antar variabel melalui data terstruktur.

Metode *Bayesian Network* digunakan karena kemampuannya dalam merepresentasikan hubungan kausal antar variabel dan menghasilkan estimasi probabilistik yang akurat. Lokasi penelitian dilaksanakan di Kota Palangka Raya, dengan melibatkan 11 perguruan tinggi dan 304 responden sivitas akademika. Data dikumpulkan menggunakan kuesioner daring berbasis *Google Form* yang disebarluaskan melalui media sosial dan grup komunikasi kampus.

Tahapan pelaksanaan penelitian meliputi: Penyusunan instrumen kuesioner berdasarkan indikator yang relevan, Penyebaran kuesioner dan pengumpulan data secara *online*, Pengolahan dan pembersihan data (*data preprocessing*), Penyusunan struktur model *Bayesian Network*, Pembelajaran parameter dan Evaluasi dan interpretasi model hasil model untuk menganalisis pola keterkaitan.

3.2 Populasi dan Sample

Populasi dalam penelitian ini adalah seluruh sivitas akademika (mahasiswa, dosen, dan tenaga kependidikan) di 11 perguruan tinggi di Kota Palangka Raya. Sampel diambil secara non-probabilistik dengan metode purposive sampling, menghasilkan 304 responden yang dianggap memenuhi kriteria partisipasi. Instrumen penelitian berupa kuesioner yang disusun berdasarkan indikator Permendikbudristek No. 30 Tahun 2021, serta dikembangkan menjadi sejumlah variabel terobservasi.

Sebaran Responden Berdasarkan Perguruan Tinggi ▼

Asal Perguruan Tinggi	Status di Perguruan Tinggi				Total
	Dosen	Mahasiswa	Orang Luar Kampus	Tenaga Kependidikan	
Institut Agama Islam Negeri Palangka Raya	0	5	1	0	6
Institut Agama Kristen Negeri Palangka Raya	1	16	0	0	17
Politeknik Kesehatan Kemenkes Palangka Raya	9	29	0	4	42
Sekolah Tinggi Ilmu Ekonomi Palangka Raya	0	1	0	0	1
Sekolah Tinggi Ilmu Hukum Tambun Bungai	0	1	0	0	1
Sekolah Tinggi Ilmu Kesehatan Eka Harap Palangka Raya	7	52	0	1	60
Sekolah Tinggi Manajemen Informatika dan Komputer Palangka Raya	0	27	0	0	27
Universitas Kristen Palangka Raya	1	4	0	0	5
Universitas Muhammadiyah Palangka Raya	0	3	0	0	3
Universitas Palangka Raya	6	125	6	3	140
Universitas Terbuka Palangka Raya	0	2	0	0	2
Total	24	265	7	8	304

Note. Each cell displays the observed counts

Gambar 1. Sebaran Responden Berdasarkan Perguruan Tinggi

3.3 Penyusunan Instrumen Kuisoner

Instrumen penelitian disusun berdasarkan indikator dari Permendikbudristek No. 30 Tahun 2021 dan referensi penelitian terdahulu. Kuesioner dirancang untuk mengumpulkan data terkait pengalaman pelecehan seksual, jenis tindakan, pelaku, lokasi, serta respons korban.

Instrumen terdiri dari 10 pertanyaan utama (Q1–Q10) yang digunakan dalam pemodelan Bayesian Network, serta pertanyaan tambahan dari SATGAS PPKS yang tidak dianalisis. Variabel target adalah Pelecehan_Seksual, hasil rekode dari Q1 melalui proses feature engineering. Semua variabel dikonversi ke bentuk kategorikal sesuai kebutuhan model probabilistik.

3.4 Pengumpulan Data

Survei dilaksanakan pada periode September 2024 hingga Februari 2025, Data dikumpulkan melalui kuesioner daring berbasis *Google Form* yang disebarluaskan melalui media sosial, grup WhatsApp, dan jaringan komunikasi kampus.

Pengambilan sampel dilakukan menggunakan metode *convenience sampling*, yaitu memilih responden yang mudah dijangkau untuk mempercepat proses pengumpulan data. Partisipasi bersifat sukarela, anonim, dan setiap responden memberikan persetujuan sebelum mengisi kuesioner.

3.5 Tahapan Pemodelan Bayesian Network

Penelitian ini menggunakan variabel-variabel hasil survei untuk membangun struktur Bayesian Network dengan bantuan perangkat lunak *GeNIe Modeler*, lalu model tersebut diterapkan untuk menganalisis faktor risiko dan memprediksi kasus pelecehan seksual di perguruan tinggi Kota Palangka Raya.

3.6 Tahapan Penelitian

Penelitian ini terdiri atas beberapa tahapan, yaitu:

a. Pra-Pemrosesan Data

Tahap pra-pemrosesan dilakukan untuk menyiapkan data sebelum proses pemodelan. Langkah-langkah yang dilakukan meliputi pembersihan data dari duplikasi dan *missing value*, penggabungan variabel berbasis *skip logic*, serta konversi seluruh variabel menjadi bentuk kategorikal. Dari proses ini diperoleh 304 data bersih dan siap digunakan.

Untuk mengatasi ketidakseimbangan pada variabel target Pelecehan_Seksual, dilakukan *random oversampling* terhadap kelas minoritas ("ya") agar distribusinya mendekati kelas mayoritas. Langkah ini penting agar model tidak bias terhadap kelas dominan dan dapat mempelajari pola dari kedua kelas secara adil.

Selanjutnya, dilakukan pembagian data menggunakan *stratified sampling* dengan rasio 80% data latih dan 20% data uji, untuk menjaga proporsi kelas tetap seimbang di kedua subset data.

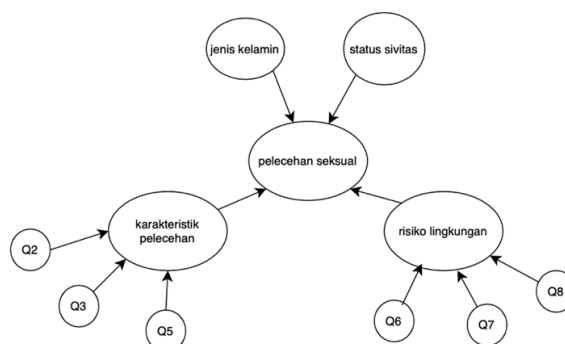
Selain itu, dilakukan *feature engineering* dengan menambahkan dua *intermediate node* untuk menyederhanakan struktur model:

- 1) Karakteristik_Pelecehan (gabungan dari Q2, Q3, Q5) diklasifikasi menjadi "tinggi" atau "rendah".
- 2) Risiko_Lingkungan (gabungan dari Q6, Q7, Q8) juga diklasifikasi menjadi dua tingkat risiko.

Penambahan *intermediate node* bertujuan menyederhanakan struktur jaringan dan mengurangi kompleksitas *Conditional Probability Table* (CPT). Dengan langkah ini, model menjadi lebih stabil, efisien, dan hasil prediksi lebih andal [14].

b. Pembuatan Struktur Jaringan

Model *Bayesian Network* dibangun menggunakan perangkat lunak *GeNIe* dengan menambahkan node untuk variabel-variabel utama seperti pelecehan_seksual, jenis_kelamin, status_pt, Q2–Q3, Q5–Q8, serta dua *intermediate node*: karakteristik_pelecehan dan risiko_lingkungan. State pada masing-masing node ditetapkan sesuai skema kategorisasi, lalu node-node tersebut dihubungkan menggunakan directed arcs berdasarkan relasi kausal.



Gambar 2. Pembuatan Struktur Jaringan

Struktur jaringan disusun secara manual dengan mempertimbangkan konteks sosial dan teori yang relevan. Pendekatan manual umum digunakan dalam kasus sosial karena mampu merepresentasikan pengetahuan domain yang tidak selalu terlihat dalam pola data.

c. Pembelajaran Parameter dan Pembagian Data

Sebelum pembelajaran parameter dilakukan, data dibagi menjadi data latih (80%) dan data uji (20%) menggunakan teknik *stratified sampling* untuk menjaga keseimbangan proporsi kelas target. Data latih digunakan untuk membangun model, sementara data uji disimpan untuk evaluasi. Pemisahan ini penting untuk menjaga objektivitas evaluasi.

Model *Bayesian Network* disusun secara manual berdasarkan pengetahuan domain, bukan menggunakan algoritma otomatis seperti PC atau K2. Struktur ini diimplementasikan dalam perangkat lunak *GeNIe Academic BayesFusion*, lalu dilakukan pembelajaran parameter menggunakan algoritma *Expectation-Maximization* (EM). Algoritma EM bekerja secara iteratif melalui dua tahap: *Expectation-step* untuk memperkirakan data yang hilang, dan *Maximization-step* untuk mengoptimalkan estimasi parameter hingga mencapai konvergensi.

Setiap node dalam jaringan kemudian memiliki Conditional Probability Table (CPT) yang menunjukkan distribusi probabilitas berdasarkan node parent-nya. Untuk mengatasi masalah *sparse data*, diterapkan parameter smoothing seperti *Laplace Correction* guna menghindari probabilitas nol.

Evaluasi awal model dilakukan dengan mengamati *log-likelihood*, sebagai ukuran seberapa baik model menjelaskan data. Setelah model terlatih, dilakukan proses *inference* menggunakan metode *AIS Sampling* (*Adaptive Importance Sampling*) untuk melakukan prediksi dan analisis skenario. Model kemudian disimpan dan digunakan sebagai alat inferensi untuk memperkirakan kemungkinan kejadian pelecehan seksual berdasarkan kondisi yang diamati.

d. Evaluasi dan Validasi

Evaluasi model dilakukan untuk mengukur kinerja *Bayesian Network* dalam memprediksi variabel target *pelecehan_seksual* berdasarkan data uji yang dipisahkan sebelumnya menggunakan *stratified sampling*. Beberapa metode evaluasi yang digunakan antara lain:

1) Akurasi Global dan Target

Evaluasi dilakukan menggunakan *maximum a posteriori* (MAP) untuk memperoleh akurasi global, serta akurasi target pada node *pelecehan_seksual* guna menilai kemampuan prediksi spesifik terhadap variabel utama [14].

2) Confusion Matrix

Confusion matrix digunakan untuk menghitung metrik evaluasi seperti *precision*, *recall*, dan *F1-score*, dengan mempertimbangkan prediksi benar dan salah pada kategori TP, TN, FP, dan FN.

3) Kurva ROC dan AUC

Evaluasi dilengkapi dengan kurva ROC dan Area *Under the Curve* (AUC), yang mengukur kemampuan model dalam membedakan kelas. Nilai AUC yang tinggi menunjukkan model bekerja baik, terutama dalam kasus dengan distribusi kelas yang tidak seimbang [15].

4) Interpretasi Hasil Evaluasi

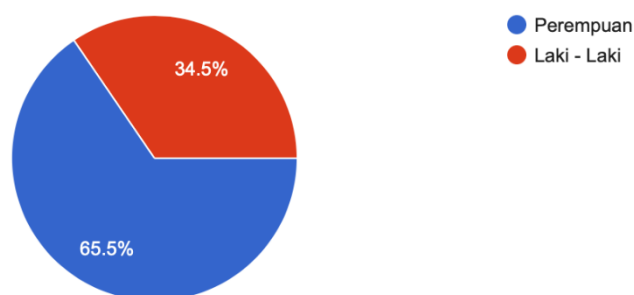
Interpretasi hasil evaluasi dilakukan untuk mengetahui sejauh mana model Bayesian Network mampu memprediksi kasus pelecehan seksual secara akurat berdasarkan data yang telah dianalisis. Hasil ini dilihat dari nilai akurasi, Confusion Matrix, ROC-AUC, dan uji validasi silang, yang secara keseluruhan menunjukkan performa model dalam membedakan antara kelas “ya” dan “tidak” pada variabel target Pelecehan_Seksual. Hasil evaluasi inilah yang akan menjadi dasar kesimpulan efektivitas model dalam melakukan prediksi.

4. PEMBAHASAN

Penelitian ini membandingkan tiga model *Bayesian Network* untuk menganalisis dan memprediksi risiko pelecehan seksual berdasarkan data survei dari sivitas akademika di Kota Palangka Raya. Model A menggunakan data asli tanpa *balancing*, Model B menggunakan *balancing* tanpa *intermediate node*, sementara Model C mengkombinasikan *balancing*, pembentukan *intermediate node*, dan pembagian data *stratified*.

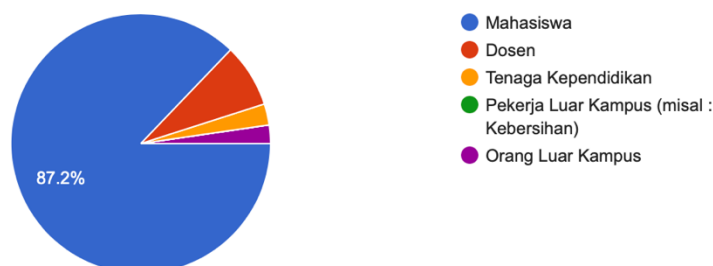
a. Data Penelitian

Dalam penelitian ini, berhasil dikumpulkan data dari 304 responden yang berasal dari berbagai kalangan sivitas akademika di beberapa perguruan tinggi di Kota Palangka Raya. Data diperoleh melalui penyebaran kuesioner online menggunakan *Google Form* selama periode September 2024 hingga Februari 2025. Responden yang mengisi terdiri dari mahasiswa, dosen, tenaga kependidikan, hingga pihak luar kampus seperti pekerja kebersihan.



Gambar 3. Diagram persentase jenis kelamin responden

Dilihat dari jenis kelamin responden, sebagian besar responden adalah perempuan, yaitu sekitar 65,5%, sedangkan laki-laki sebanyak 34,5%.



Gambar 4. Status sivitas akademika responden

Sebagian besar responden dalam penelitian ini adalah mahasiswa (87,2%), diikuti oleh dosen (8,9%), tenaga kependidikan (2,4%), serta pihak luar lainnya yang masih berkaitan dengan lingkungan kampus (masing-masing 0,7%).

b. Perbandingan Kinerja Model

Evaluasi model dilakukan menggunakan metrik akurasi, *recall*, *precision*, F1-score, dan area under the curve (AUC). Model A mencatat akurasi target tertinggi (0.980), namun memiliki kelemahan dalam mengenali korban pelecehan (*recall* hanya 78.2%). Model B menunjukkan peningkatan *recall* menjadi (80.7%) setelah penerapan *balancing* data, tetapi akurasi globalnya menurun menjadi 0.435.

Model C menunjukkan performa terbaik dalam mengenali kasus pelecehan seksual, dengan *recall* mencapai 100% dan AUC sebesar 0.863. Hal ini menunjukkan bahwa meskipun akurasi total tidak setinggi Model A, Model C lebih andal dalam mendeteksi kasus minoritas, yang sangat penting dalam konteks sosial.

1) Evaluasi Model A

$$\frac{TP + TN}{TP + TN + FP + FN} = \frac{18 + 280}{18 + 280 + 1 + 5} = 0.980 \times 100 = 98 \%$$

Gambar 5. Akurasi Model A

Nilai akurasi sebesar 98% menunjukkan bahwa Model A mampu memberikan prediksi yang benar pada sebagian besar data, baik untuk kasus positif (korban) maupun negatif (bukan korban). Namun demikian, akurat secara umum tidak selalu berarti andal dalam konteks sosial, karena model bisa saja mendeteksi hampir semua “bukan korban” dengan benar, tetapi gagal mengenali sejumlah “korban” (yang ditunjukkan oleh *recall* yang lebih rendah). Oleh karena itu, meskipun akurasinya tinggi, metrik lain seperti *recall* dan F1-score tetap harus dipertimbangkan secara bersamaan dalam evaluasi performa model.

Meskipun secara khusus akurasi pada node *pelecehan_seksual* cukup tinggi (98.0%), nilai akurasi global yang hanya 65.6% menunjukkan bahwa model belum mampu menangkap relasi secara menyeluruh antar semua node dalam jaringan. Hal ini mengindikasikan bahwa meskipun target utama terklasifikasi dengan baik, pemodelan hubungan antar variabel lainnya masih kurang optimal, sehingga bisa mempengaruhi stabilitas inferensi saat kondisi input berubah.

Dalam konteks *Bayesian Network*, akurasi global yang moderat ini dapat disebabkan oleh struktur jaringan yang belum optimal dan tidak digunakannya *intermediate node*, sehingga tabel probabilitas bersyarat (CPT) menjadi kompleks dan kurang presisi.

$$\frac{TP}{TP + FN} = \frac{18}{18 + 5} = \frac{18}{23} = 0.783 \times 100 = 78,3\%$$

Gambar 6. Nilai Recall Model A

Nilai *recall* sebesar 78.3% menunjukkan bahwa Model A mampu mendeteksi sekitar 78 dari setiap 100 korban yang sebenarnya terjadi. Meskipun ini merupakan performa yang cukup baik, masih terdapat sejumlah kasus yang tidak terdeteksi oleh model (*false negative*), yang berpotensi menimbulkan risiko pada implementasi kebijakan preventif jika model ini digunakan secara langsung di lingkungan kampus.

$$2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0.947 \times 0.947}{0.947 + 0.947} = 0.856 \times 100 = 85,6 \%$$

Gambar 7. Nilai F1-Score Model A

Nilai F1-Score sebesar 85.6% menunjukkan bahwa meskipun Model A tidak menggunakan *balancing* data maupun *intermediate node*, model ini masih mampu memberikan prediksi yang cukup stabil dan seimbang antara ketepatan (*precision*) dan sensitivitas (*recall*). Akan tetapi, *recall* sebesar 78.3% yang diperoleh pada evaluasi manual sebelumnya mengindikasikan bahwa ada kemungkinan perbedaan hasil antara evaluasi otomatis dan hitungan manual, yang bisa terjadi akibat perbedaan subset data atau parameter evaluasi yang digunakan di perangkat lunak GeNIe.

$$\frac{TP}{TP + FP} = \frac{18}{18 + 1} = \frac{18}{19} = 0.947 \times 100 = 94,7\%$$

Gambar 8. Nilai Precision Model A

Nilai *precision* sebesar 94,7% menunjukkan bahwa mayoritas prediksi “ya” oleh model memang benar-benar merupakan korban. Artinya, Model A memiliki tingkat kesalahan prediksi positif yang sangat rendah, atau dalam konteks ini, model jarang mengklasifikasikan seseorang sebagai korban secara salah (*false positive* rendah). Hal ini penting untuk menjaga kredibilitas sistem pendeteksi, terutama dalam lingkungan kampus di mana akurasi sangat krusial.

2) Evaluasi Model B

$$\frac{TP + TN}{TP + TN + FP + FN} = \frac{227 + 280}{227 + 280 + 1 + 54} = 0.902 \times 100 = 90,2 \%$$

Gambar 9. Nilai Akurasi Model B

Nilai akurasi sebesar 90,2% menunjukkan bahwa Model B memiliki kemampuan yang cukup baik secara umum dalam memprediksi data dengan benar. Namun, dibandingkan Model A, peningkatan *recall* pada Model B diperoleh dengan konsekuensi penurunan pada deteksi negatif yang akurat (lebih banyak *false negative*). Ini menunjukkan adanya *trade-off* antara akurasi umum dan sensitivitas terhadap kasus positif, yang perlu dipertimbangkan dalam konteks aplikatif seperti pemetaan risiko pelecehan seksual.

$$\frac{TP}{TP + FN} = \frac{227}{227 + 54} = \frac{227}{281} = 0.807 \times 100 = 80,7\%$$

Gambar 10. Nilai Recall Model B

Nilai *recall* ini menunjukkan bahwa model mampu mengenali 80,7% dari total kasus pelecehan seksual yang sebenarnya terjadi. Capaian ini memperlihatkan peningkatan signifikan dibanding model sebelumnya, dan menegaskan pentingnya proses balancing data untuk meningkatkan sensitivitas model terhadap kelas minoritas.

$$2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0.996 \times 0.807}{0.996 + 0.807} = 0.997 \times 100 = 99,7 \%$$

Gambar 11. Nilai F1-Score Model B

Nilai *F1-Score* sebesar 99,7% menunjukkan bahwa model memiliki kinerja yang sangat baik dalam mendeteksi kasus pelecehan seksual, dengan keseimbangan yang hampir sempurna antara ketepatan prediksi (*precision*) dan kemampuan mendeteksi semua kasus positif (*recall*). Ini menjadi indikator kuat bahwa model tidak hanya akurat secara teknis, tetapi juga dapat diandalkan secara sosial dalam konteks pencegahan dan perlindungan korban.

$$\frac{TP}{TP + FP} = \frac{227}{227 + 1} = \frac{227}{228} = 0.996 \times 100 = 99,6\%$$

Gambar 12. Nilai Precision Model B

Nilai *precision* sebesar 99,6% menunjukkan bahwa dari seluruh prediksi positif yang dilakukan oleh model, 99,6% di antaranya benar-benar merupakan kasus positif. Metrik ini sangat penting dalam konteks sosial sensitif seperti deteksi pelecehan seksual, karena *precision* yang tinggi berarti model jarang memberikan prediksi positif yang salah (*false positive*), sehingga dapat meminimalkan potensi stigma atau salah tuduh terhadap individu yang tidak terlibat.

3) Evaluasi Model C

$$\frac{TP + TN}{TP + TN + FP + FN} = \frac{56 + 35}{56 + 35 + 22 + 0} = 0.805 \times 100 = 80,5 \%$$

Gambar 13. Nilai Akurasi Model C

Nilai akurasi sebesar 80,5% menunjukkan bahwa model mampu memprediksi dengan benar 80,5% dari seluruh data uji, baik untuk kelas positif maupun negatif. Meskipun akurasi memberikan gambaran umum, dalam kasus dengan distribusi kelas yang tidak seimbang seperti pelecehan seksual, metrik ini perlu dikombinasikan dengan metrik lain seperti recall, precision, dan F1-score untuk memberikan evaluasi yang lebih menyeluruh terhadap performa model.

$$\frac{TP}{TP + FN} = \frac{56}{56 + 0} = \frac{56}{56} = 1.000 \times 100 = 100\%$$

Gambar 14. Nilai Recall Model C

Nilai *recall* sebesar 100% menunjukkan bahwa model berhasil mengidentifikasi seluruh kasus pelecehan seksual dalam data uji tanpa satu pun kesalahan deteksi. Pencapaian ini menunjukkan bahwa model sangat efektif dalam meminimalkan kesalahan tipe II (*false negative*), yang sangat krusial dalam konteks sosial yang sensitif seperti ini.

$$2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0.781 \times 1.000}{0.781 + 1.000} = 0.836 \times 100 = 83,6 \%$$

Gambar 15. Nilai F1-Score Model C

Nilai F1-Score sebesar 83,6% menunjukkan bahwa model mampu mempertahankan keseimbangan antara ketepatan prediksi (*precision*) dan kemampuan mendeteksi semua kasus positif (*recall*). Meskipun *recall* mencapai 100%, nilai *precision* yang lebih rendah menyebabkan skor harmonisasi tidak mencapai maksimum, sehingga dapat menjadi indikator perlunya peningkatan ketepatan prediksi agar tidak terjadi terlalu banyak *false positive*.

$$\frac{TP}{TP + FP} = \frac{56}{56 + 22} = \frac{56}{78} = 0.718 \times 100 = 71,8\%$$

Gambar 16. Nilai Precision Model C

Nilai *precision* sebesar 71,8% menunjukkan bahwa dari seluruh prediksi positif yang dihasilkan model, hanya 71,8% yang benar-benar merupakan kasus pelecehan seksual. Artinya, masih terdapat sejumlah *false positive* yang perlu diperhatikan karena dapat berimplikasi terhadap kemungkinan kesalahan dalam menandai individu yang sebenarnya bukan korban. Dalam konteks sosial yang sensitif, nilai *precision* yang rendah perlu diimbangi dengan *recall* tinggi untuk menjaga kehati-hatian dalam penilaian.

c. Manfaat Intermediate Node

Pembentukan *intermediate node* seperti karakteristik_pelecehan dan risiko_lingkungan dalam Model C berperan penting dalam menyederhanakan struktur jaringan dan meningkatkan keterbacaan relasi antar variabel. Node ini mengelompokkan variabel situasional dan kontekstual, sehingga memungkinkan analisis risiko yang lebih terstruktur dan logis.

Dengan adanya node ini, model dapat menangkap hubungan antar elemen lingkungan secara lebih menyeluruh. Pengelompokan variabel semacam ini membantu mengurangi jumlah parameter dalam CPT dan membuat pembelajaran parameter lebih akurat, terutama ketika data terbatas.

d. Hubungan Antar Variabel

Dalam struktur *Bayesian Network* yang dibangun, variabel target pelecehan_seksual dipengaruhi baik secara langsung maupun tidak langsung oleh sejumlah faktor. Variabel demografis seperti jenis_kelamin dan status_sivitas memiliki hubungan langsung dengan target, menunjukkan bahwa karakteristik individu turut memengaruhi risiko kejadian.

Untuk mengurangi kompleksitas dan meningkatkan interpretabilitas, variabel Q2, Q3, dan Q5 dikelompokkan dalam node perantara karakteristik_pelecehan, yang merepresentasikan jenis, frekuensi, dan konteks kejadian. Sementara variabel Q6, Q7, dan Q8 dimasukkan ke dalam node risiko_lingkungan, mencerminkan pengaruh lokasi, pelaku, dan faktor eksternal.

e. Visualisasi CPT dan Interpretasi Kombinasi Risiko

Setelah pembelajaran parameter selesai, peneliti melakukan analisis terhadap *Conditional Probability Table* (CPT) dari node pelecehan_seksual sebagai variabel target utama. CPT ini menunjukkan probabilitas seseorang mengalami pelecehan seksual berdasarkan kombinasi nilai dari empat variabel input: jenis_kelamin, status_sivitas, karakteristik_pelecehan, dan risiko_lingkungan.

jenis_kelamin	status_sivitas	karakteristik_p...	risiko_lingkungan	ya	tidak
Perempuan	Dosen	rendah	rendah	0.0625	0.9375
		tinggi	rendah	0.5	0.5
	Lainnya	rendah	tinggi	0.76190476	0.23809524
		tinggi	tinggi	0.25	0.75
Mahasiswa	Lainnya	rendah	rendah	0.94736842	0.052631579
		tinggi	rendah	0.5	0.5
	Mahasiswa	rendah	tinggi	0.0056179...	0.99438202
		tinggi	tinggi	0.5	0.5

Gambar 17. Responden Berjenis Kelamin Perempuan

jenis_kelamin	status_sivitas	karakteristik_p...	risiko_lingkungan	ya	tidak
Laki-Laki	Dosen	rendah	rendah	0.125	0.875
		tinggi	rendah	0.5	0.5
	Lainnya	rendah	tinggi	0.1	0.9
		tinggi	tinggi	0.083333333	0.91666667
Mahasiswa	Lainnya	rendah	rendah	0.16666667	0.83333333
		tinggi	rendah	0.5	0.5
	Mahasiswa	rendah	tinggi	0.013157895	0.98684211
		tinggi	tinggi	0.5	0.5

Gambar 18. Responden Berjenis Kelamin Laki-Laki

Gambar 17. dan Gambar 18. menunjukkan tampilan CPT yang dihasilkan dari aplikasi GeNIe, Dalam model *Bayesian Network* yang menggambarkan probabilitas terjadinya pelecehan seksual berdasarkan kombinasi atribut jenis kelamin, status sivitas (Dosen, Mahasiswa, Lainnya), dan risiko lingkungan (rendah/tinggi).

Tabel 1. Kombinasi dengan Nilai Probabilitas Tertinggi

Jenis kelamin	Status_sivitas	Karakteristik Pelecehan	Risiko Lingkungan	P (ya)
Perempuan	Lainnya	Tinggi	Tinggi	0.947
Perempuan	Mahasiswi	Tinggi	Tinggi	0.725
Laki-Laki	Mahasiswa	Tinggi	Tinggi	0.616
Perempuan	Dosen	Tinggi	Tinggi	0.619
Laki-Laki	Lainnya	Tinggi	Tinggi	0.5

Pada tabel 1. menemukan bahwa kombinasi risiko tertinggi pelecehan seksual terjadi pada perempuan berstatus "lainnya" dengan karakteristik pribadi dan risiko lingkungan yang tinggi ($P = 0.947$), diikuti oleh mahasiswa perempuan ($P = 0.725$) dan mahasiswa laki-laki ($P = 0.616$) dengan kondisi serupa. Sebaliknya, kombinasi dengan karakteristik dan risiko lingkungan rendah, seperti mahasiswa laki-laki dengan risiko rendah ($P = 0.013$), menunjukkan probabilitas sangat kecil. Temuan ini menegaskan bahwa status sivilitas dan lingkungan berisiko memiliki pengaruh signifikan terhadap peningkatan risiko pelecehan seksual, melampaui faktor jenis kelamin semata. Model Bayesian Network memungkinkan pemetaan interaksi antar faktor secara kuantitatif dan transparan melalui Conditional Probability Table (CPT).

Tabel 2. Kombinasi Nilai Q2, Q3, dan Q5 dengan Probabilitas Tertinggi terhadap Karakteristik_Pelecehan = "Tinggi"

Kombinasi (Q2, Q3, Q5)	Probabilitas P=(tinggi)	Keterangan
Fisik/Paksaan - Sese kali - Berduaan	0.9888	Karakteristik pelecehan berat
Fisik/Paksaan - Sese kali - Online	0.9807	Karakteristik pelecehan berat
Fisik/Paksaan - Sese kali - Berkelompok	0.9786	Karakteristik pelecehan berat
Digital - Sese kali - Online	0.0192	Karakteristik pelecehan ringan
Verbal - Sese kali - Berkelompok	0.026	Karakteristik pelecehan ringan

Pada tabel 2. menunjukkan bahwa *Analisis Conditional Probability Table* (CPT) pada *node karakteristik_pelecehan*, yang dibentuk dari variabel jenis pelecehan (Q2), frekuensi kejadian (Q3), dan situasi saat kejadian (Q5), menunjukkan bahwa kombinasi tertentu memiliki kontribusi signifikan terhadap kelas "tinggi". Probabilitas tertinggi muncul pada kombinasi pelecehan fisik atau paksaan, dengan frekuensi sese kali atau sering, dalam situasi rentan seperti berduaan atau online, misalnya $P(\text{tinggi}) = 0.9888$ (berduaan), $P = 0.9807$ (online), dan $P = 0.9786$ (berkelompok). Sebaliknya, kombinasi pelecehan digital atau verbal dengan frekuensi rendah menunjukkan probabilitas jauh lebih kecil, seperti $P = 0.0192$ (digital, online) dan $P = 0.026$ (verbal, berkelompok). Nilai probabilitas 0.5 muncul konsisten pada kombinasi yang mengandung kategori "tidak pernah", mencerminkan ketidakpastian model akibat minimnya data empiris, sehingga model bersifat netral. Dalam konteks *Bayesian Network*, nilai ini menunjukkan bahwa model tidak memiliki cukup bukti untuk menentukan kecenderungan terhadap karakteristik pelecehan yang tinggi atau rendah.

Tabel 3. Kombinasi Nilai Q6, Q7, dan Q8 dengan Probabilitas Tertinggi terhadap Risiko_Lingkungan = "Tinggi"

Kombinasi (Q2, Q3, Q5)	Probabilitas P=(tinggi)	Keterangan
Online - Mahasiswa - Alkohol/Obatan	0.984	Karakteristik pelecehan berat
Online - Dosen - Kekuasaan	0.986	Karakteristik pelecehan berat
Luar Kampus - Mahasiswa - Ancaman/Paksaan	0.964	Karakteristik pelecehan berat

Online - Mahasiswa - Lainnya	0.25	Karakteristik pelecehan ringan
Tempat Tinggal - Dosen - Lainnya	0.071	Karakteristik pelecehan ringan

Pada tabel 3. Menunjukkan bahwa *Analisis Conditional Probability Table* (CPT) pada *node risiko_lingkungan*, yang dibentuk dari variabel lokasi kejadian (Q6), status pelaku (Q7), dan faktor kekerasan (Q8), menunjukkan bahwa kombinasi tertentu memberikan kontribusi signifikan terhadap kelas “tinggi”. Risiko tertinggi ditemukan pada kombinasi lokasi online atau di luar kampus, pelaku berstatus mahasiswa atau dosen, serta faktor kekerasan seperti alkohol/obatan, ancaman/paksaan, atau kekuasaan, dengan nilai probabilitas seperti $P = 0.984$ (online, mahasiswa, alkohol), $P = 0.986$ (online, dosen, kekuasaan), dan $P = 0.964$ (luar kampus, mahasiswa, ancaman/paksaan). Sebaliknya, kombinasi yang tergolong risiko rendah umumnya melibatkan lokasi publik atau aman, pelaku dengan status tidak relevan, dan tidak adanya faktor kekerasan, seperti $P = 0.25$ (online, mahasiswa, faktor lainnya) dan $P = 0.071$ (tempat tinggal, dosen, faktor lainnya). Probabilitas netral sebesar 0.5 muncul secara konsisten pada kombinasi yang mengandung kategori “tidak pernah”, mencerminkan ketidakpastian model akibat minimnya data empiris dari responden. Dalam pendekatan Bayesian, nilai 0.5 ini menunjukkan bahwa model tidak memiliki cukup bukti untuk menentukan klasifikasi risiko, sehingga probabilitas didistribusikan secara merata.

f. Validasi Model Melalui Studi Kasus

Model C diuji menggunakan simulasi tiga studi kasus nyata yang mencerminkan skenario umum di lingkungan kampus. Hasil prediksi model mampu mengidentifikasi kasus berisiko tinggi secara konsisten, termasuk dalam konteks digital dan relasi kekuasaan. Ini menunjukkan bahwa model memiliki potensi sebagai alat bantu pengambilan kebijakan preventif.

Tabel 4. Simulasi Prediksi Faktor Risiko berdasarkan Model

Variabel	Kasus 1	Kasus 2	Kasus 3
Jenis kelamin	Perempuan	Perempuan	Perempuan
Status sivilitas	Lainnya	Mahasiswa	Mahasiswa
Q2 (jenis pelecehan)	Verbal dan Isyarat	Digital	Digital
Q3 (frekuensi)	Sesekali	Sesekali	Sesekali
Q5 (situasi)	Berkelompok	Online	Online
Q6 (lokasi)	Lingkungan_Kampus	Online	Online
Q7 (pelaku)	Lainnya	Mahasiswa	Dosen
Q8 (faktor terlibat)	Kekuasaan	Lainnya	Alkohol_Obatan
Karakteristik_Pelecehan	Tinggi	Tinggi	Tinggi
Risiko_Lingkungan	Tinggi	Tinggi	Tinggi
Prediksi Model C	Ya	Ya	Ya

g. Evaluasi Hasil Model

Hasil pemodelan dan evaluasi yang dilakukan dalam penelitian ini menunjukkan bahwa *Bayesian Network* dapat secara efektif memetakan risiko pelecehan seksual berdasarkan kombinasi variabel-variabel sosial, demografis, dan kontekstual yang tersedia dalam data. Melalui tiga pendekatan berbeda yang diuji (Model A, B, dan C), peneliti dapat mengamati perbedaan performa model berdasarkan struktur jaringan, proses balancing, serta pembentukan *intermediate node*.

Dengan demikian, peneliti menyimpulkan bahwa Model C adalah model yang paling tepat digunakan dalam konteks prediksi risiko pelecehan seksual di lingkungan kampus dengan akurasi

80.5%, Recall 100%, F1-Score 83.6% dan Precision 71.8%, karena mampu menyeimbangkan antara sensitivitas, interpretabilitas, dan keterwakilan data minoritas.

5. KESIMPULAN

Pemodelan menggunakan *Bayesian Network* efektif dalam menganalisis faktor risiko dan memprediksi kejadian pelecehan seksual di lingkungan perguruan tinggi. Hasil analisis *Conditional Probability Table* (CPT) mengungkap bahwa kombinasi karakteristik kejadian berat, lingkungan berisiko tinggi, dan faktor demografis tertentu (seperti jenis kelamin perempuan dan status sivitas non-mahasiswa) secara signifikan meningkatkan probabilitas terjadinya pelecehan.

Pembentukan *intermediate node* seperti *karakteristik_pelecehan* dan *risiko_lingkungan* terbukti menyederhanakan struktur jaringan dan meningkatkan akurasi inferensi. Model terbaik (Model C) mampu mengenali seluruh kasus korban dengan recall mencapai 100%, serta menghasilkan struktur yang efisien untuk prediksi berbasis data.

Dengan performa yang tinggi dan interpretasi yang kuat, model ini dapat dimanfaatkan sebagai alat bantu dalam pemetaan risiko serta perumusan strategi pencegahan pelecehan seksual yang lebih terarah di lingkungan kampus.

6. DAFTAR PUSTAKA

- [1] A. Tadei, J. Pensar, J. Corander, K. Finnilä, P. Santtila, and J. Antfolk, "A Bayesian Decision-Support Tool for Child Sexual Abuse Assessment and Investigation," *Sexual Abuse: Journal of Research and Treatment*, vol. 31, no. 4, pp. 374–396, Jun. 2019, doi: 10.1177/1079063217732791.
- [2] C. Bradshaw and D. M. Blei, "A BAYESIAN MODEL OF UNDERREPORTING FOR SEXUAL ASSAULT ON COLLEGE CAMPUSES." [Online]. Available: <https://nces.ed.gov/ipeds/>
- [3] "PROBABILISTIC REASONING IN INTELLIGENT SYSTEMS: Networks of Plausible Inference REVISED SECOND PRINTING."
- [4] R. J. McNally, A. Heeren, and D. J. Robinaugh, "A Bayesian network analysis of posttraumatic stress disorder symptoms in adults reporting childhood sexual abuse," *Eur J Psychotraumatol*, vol. 8, no. sup3, 2017, doi: 10.1080/20008198.2017.1341276.
- [5] Sulandjari, "Pelecehan Seksual dan Budaya Patriarki," *Jurnal Sosiologi Reflektif*, vol. 11, no. 2, pp. 187–202, 2017.
- [6] M. Arif, "Relasi Kuasa dalam Kasus Kekerasan Seksual di Lingkungan Pendidikan Tinggi," *Jurnal Hukum & Gender*, vol. 5, no. 1, pp. 45–60, 2017.
- [7] Komnas Perempuan dan Koalisi Ruang Publik Aman, *Laporan Survei Nasional Kekerasan Seksual di Lingkungan Kampus*, Jakarta, 2020.
- [8] N. Hikmah, "Pentingnya Penanganan Institusional dalam Kasus Kekerasan Seksual di Perguruan Tinggi," *Jurnal Pendidikan dan Perlindungan Anak*, vol. 3, no. 2, pp. 72–80, 2017.
- [9] Annisa dan Hendro, "Analisis Faktor Penyebab Pelecehan Seksual dalam Perspektif Gender," *Jurnal Ilmu Sosial dan Humaniora*, vol. 3, no. 1, pp. 55–62, 2014.
- [10] R. K. Merton, *Social Theory and Social Structure*, New York: The Free Press, 1938.
- [11] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, San Mateo, CA: Morgan Kaufmann, 1988.
- [12] K. B. Korb and A. E. Nicholson, *Bayesian Artificial Intelligence*, 2nd ed., Boca Raton, FL: Chapman and Hall/CRC, 2010.
- [13] F. V. Jensen and T. D. Nielsen, *Bayesian Networks and Decision Graphs*, 2nd ed., New York: Springer, 2007.
- [14] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*, Cambridge, MA: MIT Press, 2009.
- [15] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.