



Hak Cipta Dilindungi Undang-Undang

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia. Plagiarisme juga dilarang dan dapat dikenakan sanksi.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

Universitas Islam Indragiri

ANALISIS SENTIMEN ULASAN APLIKASI JOBSTREET PADA GOOGLE PLAY STORE MENGGUNAKAN SMOTE DENGAN ALGORITMA XGBOOST DAN RANDOM FOREST

SKRIPSI



DISUSUN OLEH :

MUHAMAD ADIF

403221010072

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN ILMU KOMPUTER
UNIVERSITAS ISLAM INDRAGIRI
2026**

PERSETUJUAN

ANALISIS SENTIMEN ULASAN APLIKASI JOBSTREET PADA GOOGLE PLAY STORE MENGGUNAKAN SMOTE DENGAN ALGORITMA XGBOOST DAN RANDOM FOREST

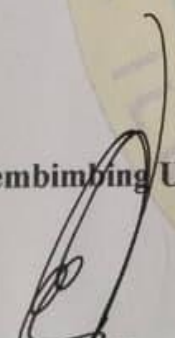
Yang dipersiapkan dan disusun oleh

MUHAMAD ADIF

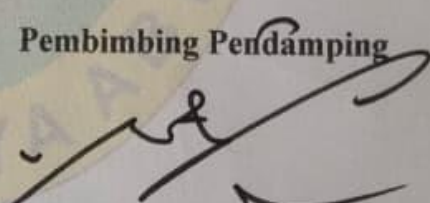
403221010072

Telah disetujui oleh Dosen Pembimbing
pada tanggal 15 Januari 2026

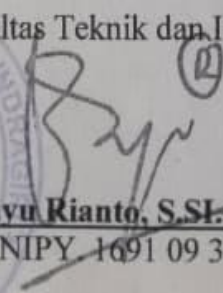
Pembimbing Utama


Samsudin, S.Kom, M.Kom
NIDN. 1009098501

Pembimbing Pendamping


Muh. Rasyid Ridha, S.SI., M.Kom
NIDN. 1013089001

Dekan Fakultas Teknik dan Ilmu Komputer


Dr. Bayu Rianto, S.SI., M.Kom
NIPY 1691 09 321

PENGESAHAN

ANALISIS SENTIMEN ULASAN APLIKASI JOBSTREET PADA GOOGLE PLAY STORE MENGGUNAKAN SMOTE DENGAN ALGORITMA XGBOOST DAN RANDOM FOREST

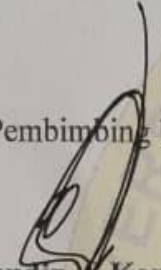
Dipersiapkan dan disusun oleh

MUHAMAD ADIF
403221010072

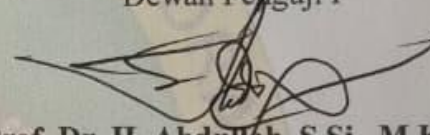
Telah Diuji dan Dipertahankan di depan Dewan Penguji
pada Tanggal 24 Juli 2026

Susunan Dewan Penguji

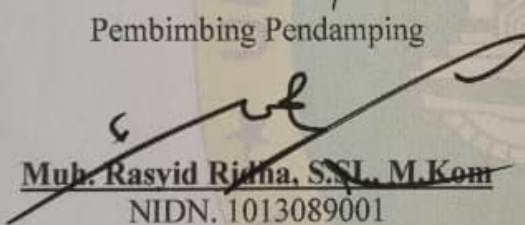
Pembimbing Utama


Samsudin, S.Kom, M.Kom
NIDN. 1009098501

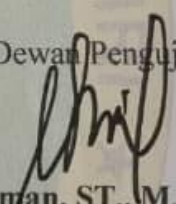
Dewan Penguji I


Prof. Dr. H. Abdullah, S.Si., M.Kom
NIDN. 1008037001

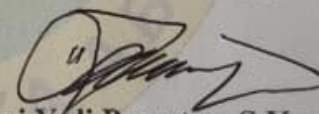
Pembimbing Pendamping


Muh. Rasyid Ridha, S.Si., M.Kom
NIDN. 1013089001

Dewan Penguji II

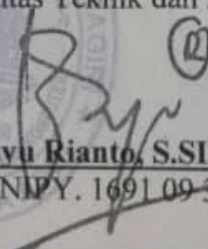

Usman, ST., M.Kom
NIDN. 1017078302

Dewan Penguji III

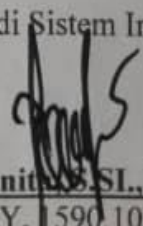

Dwi Yuli Prasetyo, S.Kom., M.Kom
NIDN. 1020078602

Tugas Akhir ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer

Dekan Fakultas Teknik dan Ilmu Komputer


Dr. Bayu Rianto, S.Si., M.Kom
NIPY. 1691 09 321

Ka. Prodi Sistem Informasi


Fitri Yunita, S.Si., M.Kom
NIPY. 1590 10 293

SURAT PERNYATAAN

Yang bertanda tangan dibawah ini :

Nama Mahasiswa : **MUHAMAD ADIF**

NIM : 403221010072

Program Studi : Sistem Informasi

Menyatakan bahwa skripsi sbb:

Judul bahasa Indonesia :

ANALISIS SENTIMEN ULASAN APLIKASI JOBSTREET PADA GOOGLE PLAY STORE MENGGUNAKAN SMOTE DENGAN ALGORITMA XGBOOST DAN RANDOM FOREST

Judul bahasa Inggris :

Sentiment Analysis of JobStreet App Reviews on the Google Play Store Using SMOTE with XGBoost and Random Forest Algorithms

Pembimbing Utama : Samsudin, S.Kom, M.Kom

Pembimbing Pendamping : Muh. Rasyid Ridha, S.SI., M.Kom

1. Laporan Tugas Akhir ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas Islam Indragiri maupun di perguruan tinggi lainnya;
2. Laporan Tugas Akhir ini merupakan gagasan, rumusan dan pemikiran SAYA sendiri, tanpa bantuan pihak lain kecuali arahan Tim Dosen Pembimbing;
3. Dalam laporan Tugas Akhir ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada laporan ini;
4. Perangkat lunak yang digunakan dalam Tugas Akhir ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas Islam Indragiri;
5. Pernyataan ini SAYA buat dengan sesungguhnya apabila di kemudian hari terdapat penyimpangan dan ketidak benaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku diperguruan tinggi.

Tembilahan, 24 Juli 2026

atakan,



MUHAMAD ADIF



KATA PENGANTAR

“Data bukan sekadar angka, tetapi jejak cerita manusia yang menunggu untuk dipahami.”

Puji syukur ke hadirat Allah Swt, tuhan yang maha Esa atas rahmat, kasih, dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Sentimen Ulasan Aplikasi JobStreet di Google Play Store Menggunakan SMOTE dengan Algoritma XGBoost dan Random Forest”.

Perjalanan saat menyusun skripsi ini bukanlah hal yang mudah. Berbagai tantangan, keraguan, dan keterbatasan sempat menjadi bagian dari proses yang harus dilalui. Namun, di balik setiap proses penyusunan selalu ada dukungan, doa, dan semangat dari orang-orang terdekat yang menjadi energi bagi penulis untuk terus melangkah hingga sampai sekarang ini. Skripsi ini bukan hanya sekadar karya ilmiah, tetapi juga menjadi bukti dari proses belajar, perjuangan, dan ketekunan yang telah dilalui.

Tanpa mengurangi rasa hormat penulis berterima kasih kepada rekan-rekan yang terlibat dalam penyusunan skripsi, yaitu,:

1. Bapak Dr. H. Najamuddin, Lc., M.A, sebagai Rektor Universitas Islam Indragiri.
2. Dr. Bayu Rianto, S.Si., M.,Kom selaku Dekan Fakultas Teknik dan Ilmu Komputer Universitas Islam Indragiri,.
3. Ibu Fitri Yunita, S.SI., M.Kom, selaku Ketua Program Studi Sistem Informasi.
4. Bapak Samsudin, S.Kom., M.Kom., selaku pembimbing pertama yang berjasa membimbing, mengarahkan dan memberi masukan dengan sabar dalam penulisan skripisi ini.



5. Bapak Muh Rasyid Ridha, S.SI., M.Kom., selaku pembimbing penulisan skripsi ini.
6. Para dosen penguji yang telah meluangkan waktu, memberikan kritik, saran, serta masukan yang sangat berarti demi penyempurnaan skripsi ini.
7. Bapak Sudirman dan Ibu Eliani selaku kedua orangtua yang sangat saya cintai, inti dari segala perjuangan saya, dan menjadi alasan saya mengapa saya sampai ke titik ini, sekali lagi terima kasih banyak atas ketulusan, kesabaran serta dukungannya selama ini yang tidak akan mampu semuanya diutarakan oleh kata.
8. Semua pihak yang terlibat membantu secara langsung maupun tidak, tidak bisa saya sebutkan satu persatu.
9. Terakhir, penulis mempersembahkan skripsi ini spesial untuk orang yang selalu bertanya kapan kamu wisuda? Dan kapan skripsimu selesai? Wisuda hanya seremonial akhir setelah melewati berbagai proses, terlambat lulus atau tidak, lulus tepat waktu bukanlah sebuah kejahatan dan bukanlah sebuah aib. Alangkah kerdilnya jika kecerdasan seseorang diukur dari cepatnya wisudanya. Bukankah sebaik-baiknya skripsi adalah skripsi yang selesai entah itu tepat waktu maupun tidak.

Penulis menyadari skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis dengan rendah hati menerima segala kritik dan saran yang membangun sebagai bahan perbaikan di masa yang akan datang.

Akhir kata, penulis berharap semoga skripsi ini dapat memberikan manfaat, tidak hanya bagi pengembangan ilmu pengetahuan, tetapi juga sebagai langkah

kecil yang dapat memberikan kontribusi nyata, khususnya dalam pemanfaatan teknologi untuk memahami kebutuhan dan pengalaman pengguna.

Tembilahan, 24 Juli 2026

Penulis

Muhamad Adif



Hak Cipta Dilindungi Undang-Undang

Universitas Islam Indragiri

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

DAFTAR ISI



Hak Cipta Dilindungi Undang-Undang

Universitas Islam Indragiri

HALAMAN JUDUL	
PERSETUJUAN	
PENGESAHAN	
PERNYATAAN	
KATA PENGANTAR	v
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
INTISARI	xii
ABSTRACT	xiii
BAB I PENDAHULUAN	
1.1 Latar Belakang.....	
1.2 Rumusan Masalah.....	
1.3 Batasan Penelitian.....	
1.4 Tujuan Penelitian	
1.5 Manfaat Penelitian	
1.6 Sistematika Penelitian.....	
BAB II TINJAUAN LITERATUR	
2.1 Analisis Sentimen	
2.2 Text Mining	
2.3 Pelabelan Sentimen.....	
2.4 Preprocessing Teks	
2.5 TF-IDF (Term Frequency-Inverse Document Frequency)	
2.6 SMOTE.....	
2.7 XGBoost	
2.8 Random Forest.....	
2.9 Confusion Matrix.....	
2.10 Evaluasi Model	
2.11 Penelitian Terdahulu	
2.12 Rangkuman.....	
BAB III METODE PENELITIAN	
3.1 Jenis Penelitian	



Hak Cipta Dilindungi Undang-Undang

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.
Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.
Plagiarisme juga dilarang dan dapat dikenakan sanksi.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

Universitas Islam Indragiri

3.2	Kerangka Penelitian
3.3	Teknik Pengumpulan Data.....
3.4	Seleksi Data
3.5	Pelabelan Sentimen.....
3.6	Preprocessing Teks
3.7	Ekstraksi Fitur Menggunakan TF-IDF
3.8	Pembagian Data Training dan Testing
3.9	Penerapan SMOTE
3.10	Klasifikasi Menggunakan XGBoost.....
3.11	Klasifikasi Menggunakan Random Forest
3.12	Evaluasi Model
3.13	Perbandingan Hasil Algoritma.....
3.14	Perangkat dan Tools Penelitian.....
BAB IV HASIL PENELITIAN	
4.1	Hasil Pengumpulan data
4.2	Hasil Seleksi Data
4.3	Hasil Pelabelan Sentimen.....
4.4	Hasil Preprocessing Data
4.4.1	Cleaning
4.4.2	Case Folding.....
4.4.3	Tokenizing.....
4.4.4	Stop Removal
4.4.5	Stemming
4.5	Hasil Ekstraksi Fitur TF-IDF
4.6	Hasil Pembagian Data.....
4.7	Hasil Penerapan SMOTE
4.8	Hasil Klasifikasi XGBoost.....
4.9	Hasil Klasifikasi Random Forest
4.10	Analisis Performa Algoritma
4.11	Pembahasan Hasil Sentimen
BAB V KESIMPULAN DAN SARAN	
5.1	Kesimpulan.....
5.2	Saran.....
DAFTAR PUSTAKA	



1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	
Tabel 3. 1 Format data awal	
Tabel 3. 2 Tahapan Seleksi Data	
Tabel 3. 3 Pelabelan Awal	
Tabel 3. 4 Tahapan Proprocessing	
Tabel 3. 5 Skenario Pembagian Data Training dan Testing	
Tabel 3. 6 Distribusi Data Sebelum dan Sesudah Penerapan SMOTE ...	
Tabel 3. 7 Tahapan Klasifikasi Menggunakan XGBoost ..	
Tabel 3. 8 Tahapan Klasifikasi Menggunakan Random Forest	
Tabel 3. 9 Metrik Evaluasi Model.....	
Tabel 3. 10 Rencana Perbandingan Hasil Evaluasi XGBoost dan Random Forest...	
Tabel 3. 11 Perangkat Keras	
Tabel 3. 12 Perangkat Lunak.....	
Tabel 4. 1 Pembagian Data	
Tabel 4. 2 Perbandingan Performa Algoritma.....	



Hak Cipta Dilindungi Undang-Undang

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.
Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi.
Plagiarisme juga dilarang dan dapat dikenakan sanksi.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

Universitas Islam Indragiri

DAFTAR GAMBAR

Gambar 3. 1 Kerangka Penelitian	
Gambar 4. 1 Data Ulasan Hasil Pengumpulan Data	
Gambar 4. 2 Tampilan Data Ulasan Aplikasi Jobstreet	
Gambar 4. 3 Hasil Seleksi Data	
Gambar 4. 4 Hasil Pelabelan Sentimen.....	
Gambar 4. 5 Diagram Distribusi Ulasan Aplikasi Jobstreet	
Gambar 4. 6 Hasil Tahap Cleaning	
Gambar 4. 7 Hasil Tahap Case Folding	
Gambar 4. 8 Hasil Tahap Tokenizing	
Gambar 4. 9 Hasil Tahap Stopword Removal	
Gambar 4. 10 Hasil Tahap Stemming dan Hasil Akhir Preprocessing	
Gambar 4. 11 Hasil Matriks TF-IDF	
Gambar 4. 12 Hasil Ekstraksi Fitur TF-IDF	
Gambar 4. 13 Hasil Distribusi Data Training	
Gambar 4. 14 Distribusi Data Training Sebelum dan Sesudah SMOTE	
Gambar 4. 15 Confusion Matrix XGBoost	
Gambar 4. 16 Confusion Matrix XGBoost	
Gambar 4. 17 Hasil Classification Report Algoritma XGBoost .	
Gambar 4. 18 Confusion Matrix Random Forest...	
Gambar 4. 19 Confusion Matrix Random Forest...	
Gambar 4. 20 Classification Report Random Forest	
Gambar 4. 21 Grafik Perbandingan Performa Algoritma XGBoost dan Random Forest.....	
Gambar 4. 22 WordCloud Sentimen Positif	
Gambar 4. 23 WordCloud Sentimen Negatif.....	



1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

INTISARI

Penelitian ini bertujuan menganalisis sentimen ulasan pengguna aplikasi JobStreet pada Google Play Store dengan menerapkan teknik oversampling SMOTE serta membandingkan performa algoritma XGBoost dan *Random Forest*. Penelitian ini dilatarbelakangi oleh tingginya jumlah ulasan pengguna yang menghasilkan data teks tidak terstruktur, serta adanya ketidakseimbangan jumlah data pada kelas sentimen yang berpotensi menurunkan kinerja model klasifikasi. Data penelitian diperoleh melalui proses *scraping* sebanyak 9.000 ulasan, kemudian setelah tahap seleksi diperoleh 6.510 data yang digunakan dalam proses analisis. Pelabelan sentimen dilakukan berdasarkan rating pengguna, kemudian data diproses melalui tahapan *preprocessing*. Selanjutnya, data diekstraksi menggunakan metode TF-IDF, dibagi menjadi data latih dan data uji dengan rasio 80:20, serta diterapkan teknik SMOTE pada data latih untuk mengatasi ketidakseimbangan kelas. Hasil penelitian menunjukkan bahwa algoritma XGBoost menghasilkan nilai *accuracy* sebesar 84,79% dengan *F1-score* sebesar 0,848, sedangkan algoritma *Random Forest* memperoleh *accuracy* sebesar 84,56% dengan *F1-score* sebesar 0,846. Berdasarkan hasil tersebut, XGBoost menunjukkan performa yang sedikit lebih baik dibandingkan *Random Forest* dalam klasifikasi sentimen ulasan aplikasi JobStreet. Selain itu, hasil analisis menunjukkan bahwa sentimen positif didominasi oleh kemudahan dalam mencari lowongan pekerjaan,



sedangkan sentimen negatif lebih banyak berkaitan dengan kendala teknis, seperti masalah login dan kesulitan dalam penggunaan aplikasi.

Kata Kunci: Analisis Sentimen, JobStreet, SMOTE, XGBoost, *Random Forest*.

ABSTRACT

This study aims to analyze the sentiment of JobStreet application user reviews on Google Play Store by applying the Synthetic Minority Over-sampling Technique (SMOTE) and comparing the performance of the XGBoost and Random Forest algorithms. The study is motivated by the large number of user reviews, which generate unstructured textual data, as well as the imbalance between sentiment classes that may reduce the performance of classification models. The data were collected through a web scraping process, resulting in 9,000 reviews. After the data selection stage, 6,510 reviews were retained for analysis. Sentiment labels were assigned based on user ratings, and the data then underwent several preprocessing. Furthermore, the text data were transformed into numerical features using the TF-IDF method. The dataset was divided into training and testing sets with an 80:20 ratio, and SMOTE was applied to the training data to address class imbalance. The experimental results show that the XGBoost algorithm achieved an accuracy of 84.79% with an F1-score of 0.848, while the Random Forest algorithm achieved an accuracy of 84.56% with an F1-score of 0.846. These findings indicate that XGBoost performed slightly better than Random Forest in classifying the sentiment of JobStreet application reviews. In addition, positive reviews were generally associated with the ease of searching for job opportunities, whereas

negative reviews mainly highlighted technical issues, such as login problems and application usability.

Keywords: Sentiment Analysis, JobStreet, SMOTE, XGBoost, Random Forest.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

Universitas Islam Indragiri

