

APBD ANOMALY DETECTION

FINANCE AND HEALTH HACKATHON 2022 : HEALTH AND FINANCE DATA ANALYTICS

by : JLS



Transparency International Indonesia :

"Uang rakyat dalam praktek APBD menguap sekitar 30-40 persen oleh perilaku korupsi."

Total Estimasi:

Rp. 341 – 455 Triliun

APBD adalah salah satu objek yang paling dikorupsi oleh kepala daerah di Indonesia di tahun 2017 .¹



1. <https://antikorupsi.org/id/article/tren-korupsi-2017-objek-penyalahgunaan-apbd-paling-banyak-dikorupsi-oleh-kepala-daerah>
2. <https://news.detik.com/berita/d-5848377/bamsoet-sebut-40-apbnapbd-menguap-oleh-korupsi>
3. <https://djpk.kemenkeu.go.id/portal/data/apbd?tahun=2021&provinsi=--&pemda=-->

Our Team



Joshia Cahyadi

Mathematics Student at
Institut Teknologi Bandung



Steven Alexander Liong

Mathematics Student at
Institut Teknologi Bandung



Leonardo Ignatius

Computer Science Student at
Bina Nusantara University

Problems, Solutions, and Goals

Problem

Penyalahgunaan APBD adalah salah satu objek yang **paling sering dikorupsi** oleh kepala daerah di Indonesia tahun 2017

Solutions

Kami mengusulkan pembuatan model **Pendeteksi Anomaly APBD** yang dibuat dengan **Machine Learning**

Goals

Untuk menemukan anomali pada RAPBD sehingga dapat ditinjau ulang dan **mengurangi kemungkinan korupsi APBD**.

Data

- Kami menggunakan File **A2021_dataset contoh.csv** untuk data APBD Tahun 2021.
- Kolom yang digunakan :**
 - kodesatker*
 - namapemda*
 - statusdata*
 - namaurusanprogram*
 - standarobjek*
 - namaakunobjek*
 - namaskpd*
 - namaprogram*
 - namakegiatan*
 - namafungsi*
 - standarfungsimappp*
 - namaakunrincian*
 - namaakunsub*
 - nilaianggaran*
 - prov*
- 70%** dari penyalahgunaan APBD misused terdapat pada **belanja barang dan jasa** oleh pemerintah.
- Akun Jenis* yang digunakan adalah **belanja barang dan jasa**, dimana akun jenis ini meliputi **82% dari seluruh data**.

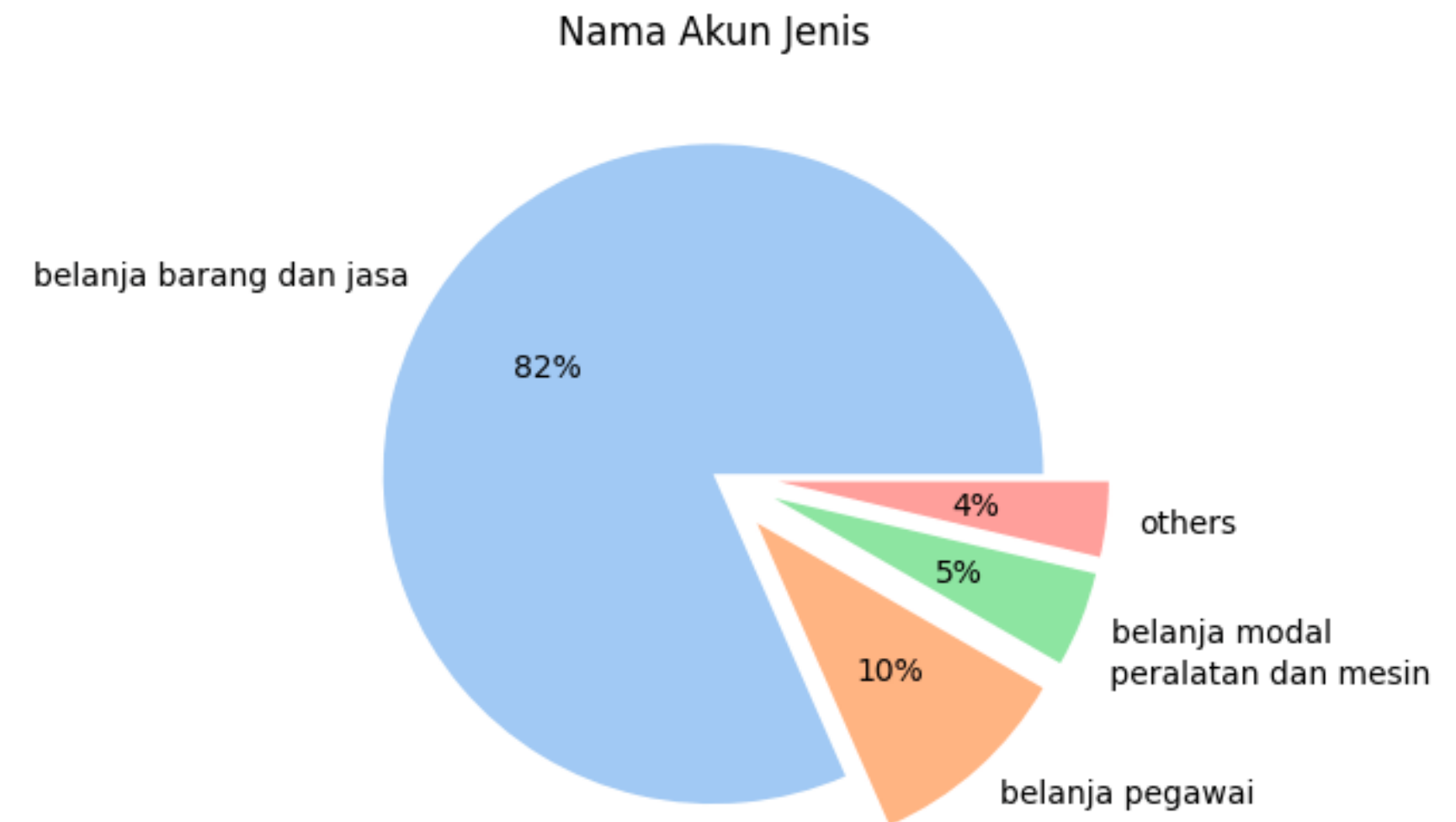
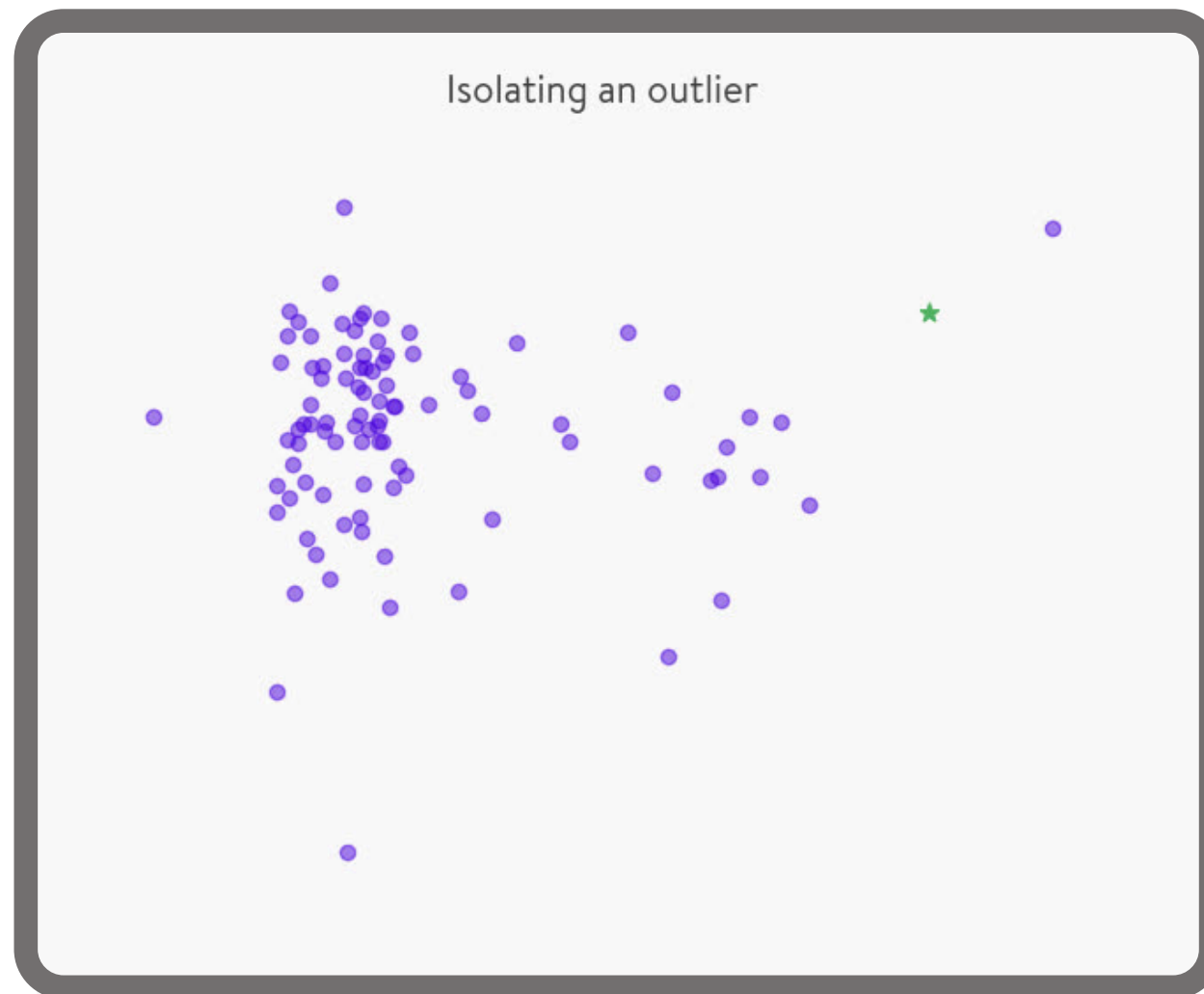


Figure 1. Proportion of *Nama Akun Jenis*'s types

Method

Isolation Forest



FGambar 2. Ilustrasi Isolation Forest Algorithm

Apa itu isolation forest algorithm?

Isolation forest adalah algoritma yang berdasarkan Decision Tree. Dimana isolation forest memiliki asumsi bahwa data yang **sedikit dan berbeda** adalah data yang **anomali**.

Bagaimana Cara Kerjanya?

Data **Anomali** akan memiliki **partisi yang lebih sedikit** untuk terisolasi. sebagai ilustrasi, panjang tree yang dilewati data anomali lebih pendek dibandingkan data yang normal

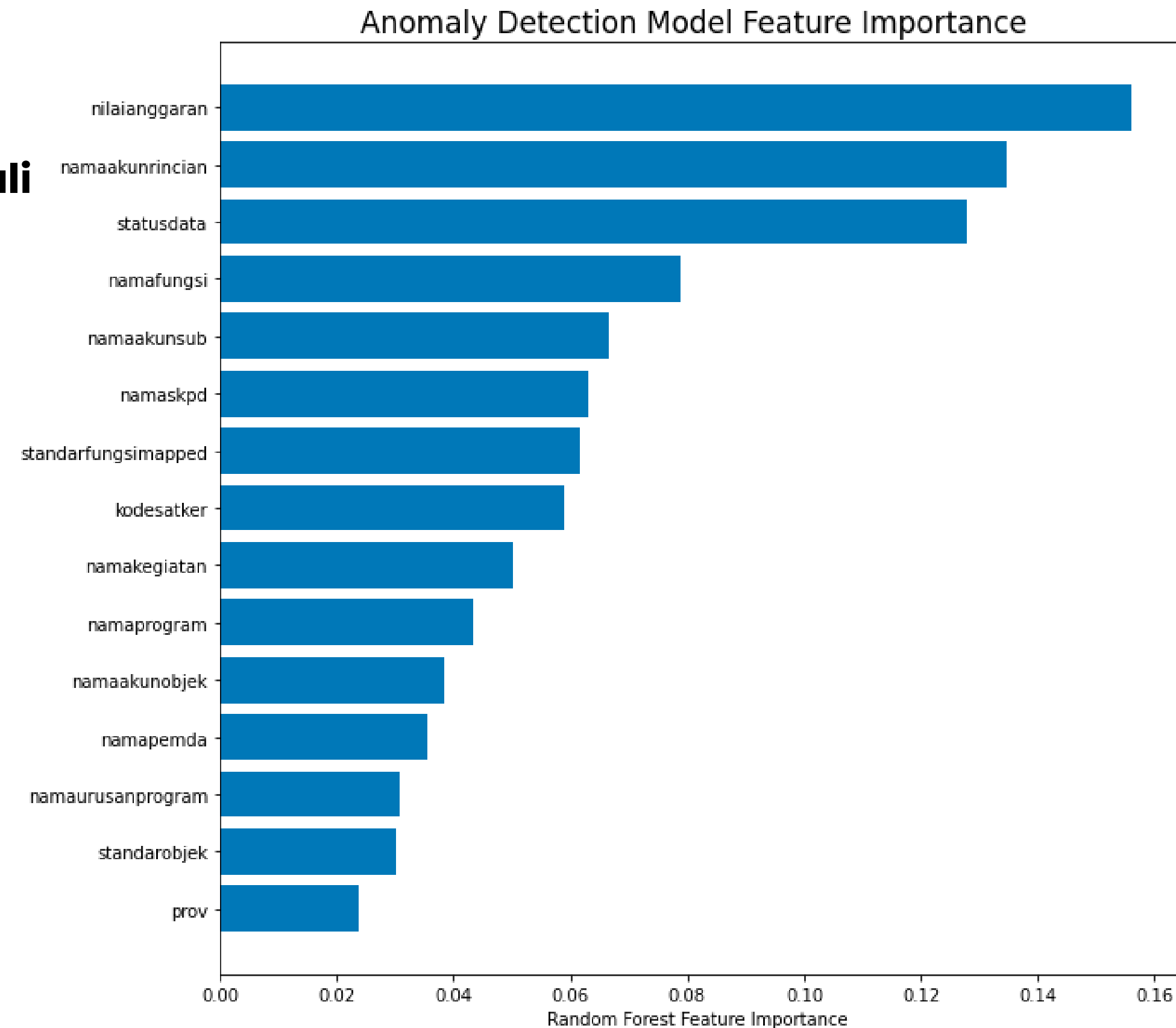
Result

Fitur-fitur yang penyebab **Kemungkinan Anomali**

Kami menggunakan **algoritma random forest** untuk menentukan **feature importance** untuk setiap fitur dalam dataset. Hasil yang kami dapat :

- ***nilaianggaran***
- ***namaakunrincian***
- ***statusdata***
- ***namafungsi***
- ***namaakunsub***

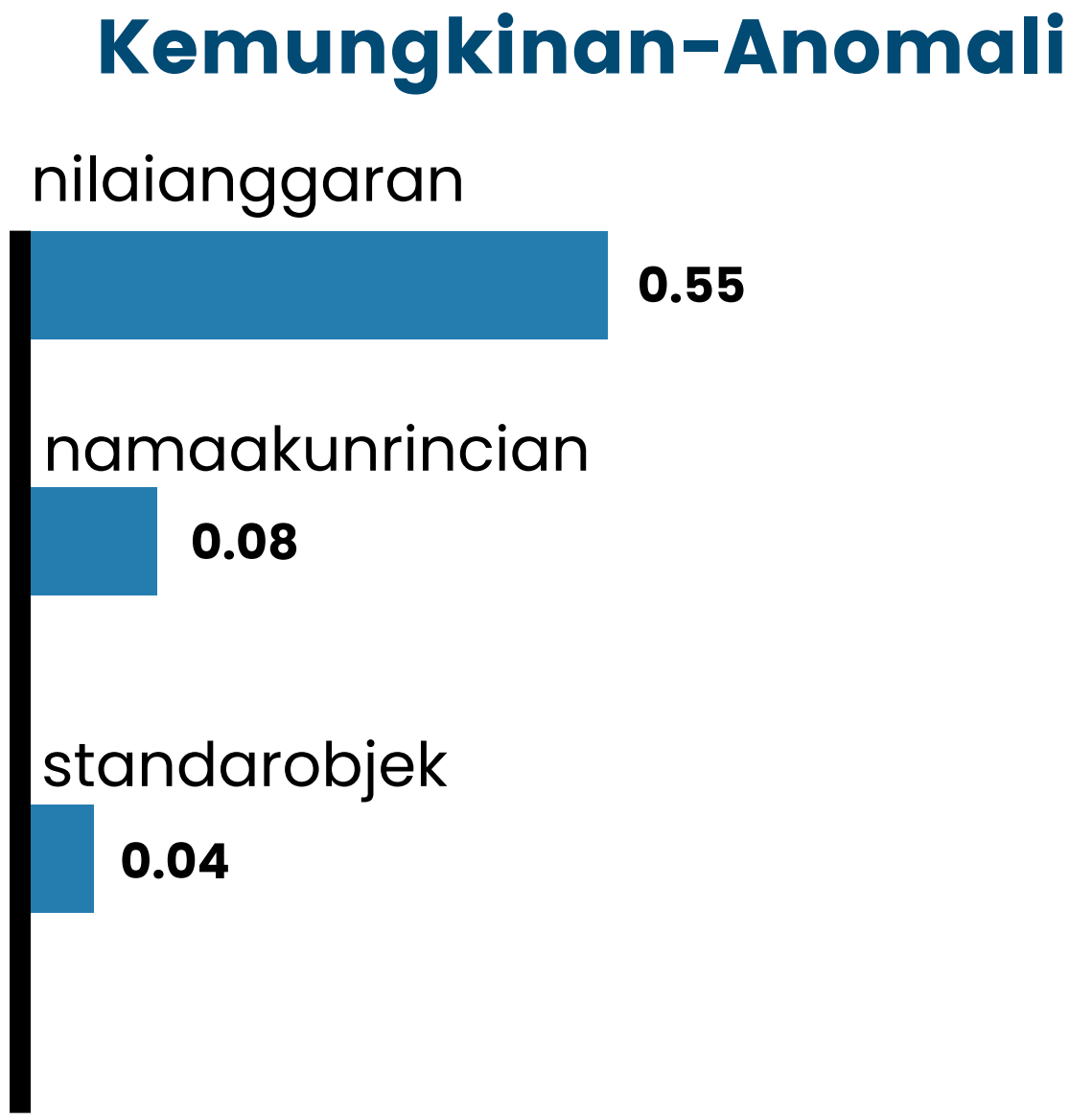
as sebagai **fitur-fitur yang memiliki pengaruh paling besar** dalam mendeteksi kemungkinan anomali.



Gambar 3. The Level of Importance Untuk Setiap Fitur

Result

Contoh Kasus **kemungkinan-anomali**

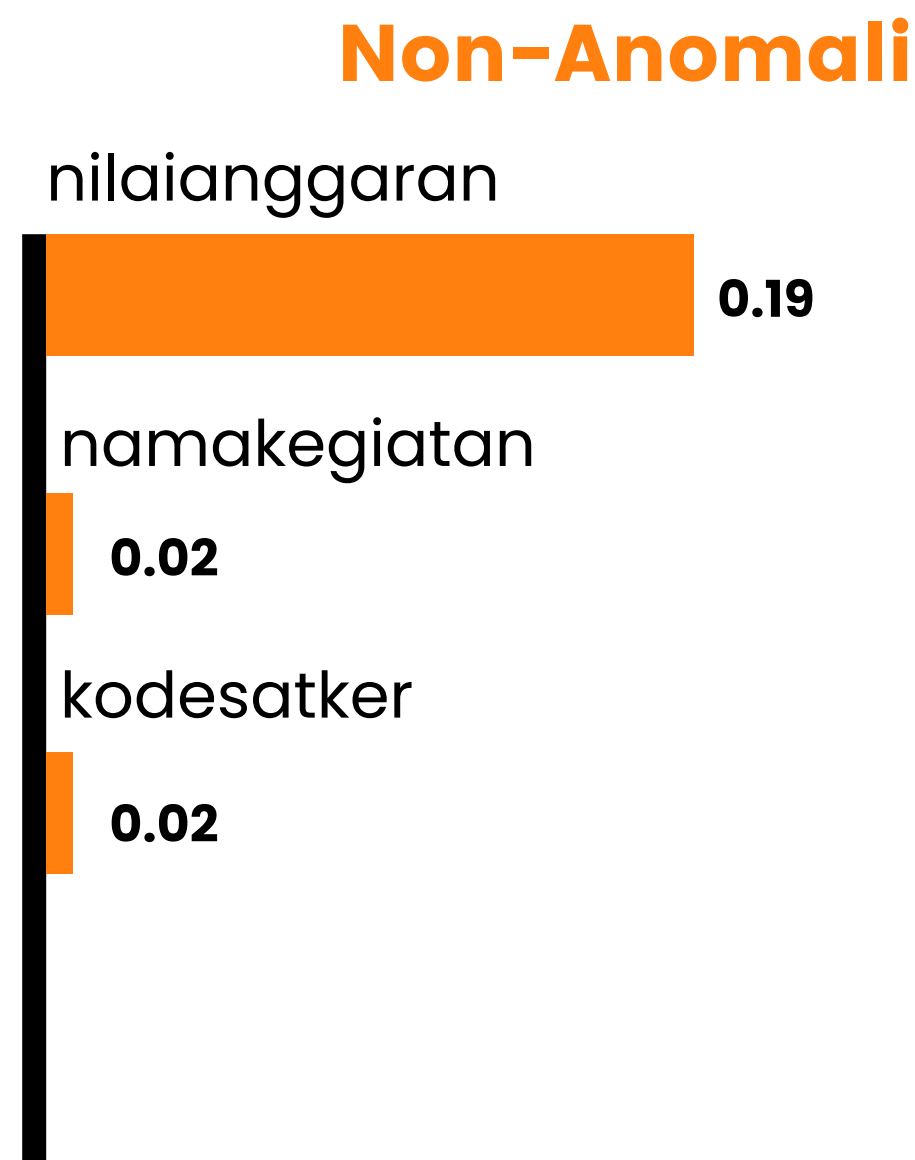


Feature	Value
nilaianggaran	2270930000.00
namaakunrincian	639.00
standarobjek	18.00
namaakunobjek	0.00
namaskpd	898.00
kodesatker	990881.00
prov	10.00
namaurusanprogram	348.00
namafungsi	484.00
namaakunsub	2335.00

Gambar 3. Fitur pada Kasus Contoh Kemungkinan-Anomali

Result

Contoh Kasus **Non-anomali**



Feature	Value
nilaianggaran	11125000.00
namaakunrincian	639.00
namaakunobjek	29.00
standarobjek	18.00
namakegiatan	17257.00
namafungsi	619.00
kodesatker	991066.00
namaurusanprogram	284.00
prov	10.00
namapemda	170.00

Gambar 4. Fitur pada Kasus Contoh Non-Anomali

Conclusion

Berdasarkan proyek yang kami buat, terdapat beberapa **poin penting**:

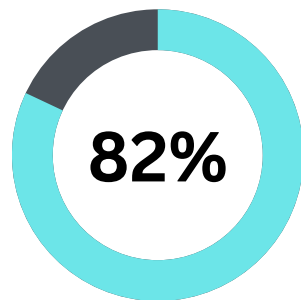
1

Objek yang paling dikorupsi dan Akun Jenis yang paling banyak digunakan



Uang rakyat dalam praktek APBD **menguap** sekitar **30-40 persen** oleh perilaku **korupsi** ~ **Rp. 341 - 455 Triliun**

Akun Jenis yang Digunakan:



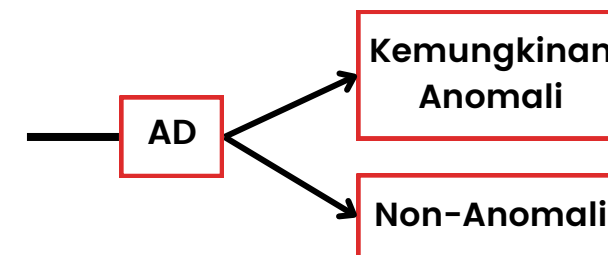
of all data

Belanja barang dan jasa (82%)

Akun jenis dengan tingkat **korupsi paling tinggi**, yaitu **70 persen**.

2

Kami membangun **model detektor anomali** untuk Akun Jenis **Belanja Barang dan Jasa**, untuk **mendeteksi kemungkinan anomali**



Dengan input **rancangan APBD**, model detektor anomali akan menghasilkan satu keluaran; **kemungkinan-anomali** dan **non-anomali**

Dan fitur yang memiliki pengaruh paling besar adalah :

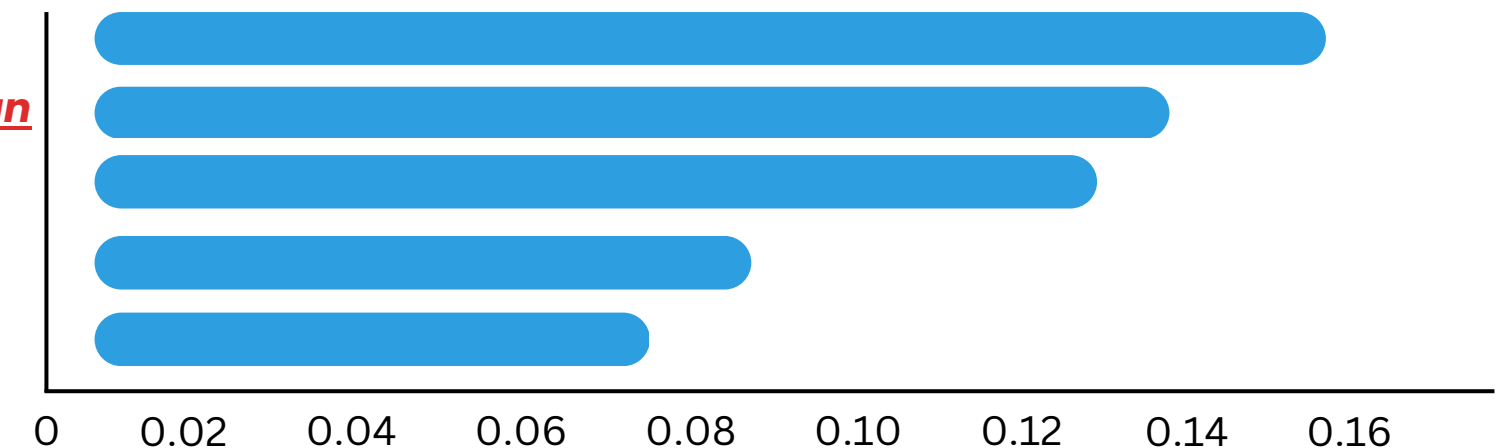
nilaianggaran

namaakunrincian

statusdata

namafungsi

namaakunsub



Kami menyimpulkan bahwa model ini dapat **mendeteksi kemungkinan anomali** pada data rancangan APBD, yang lebih lanjutnya dapat digunakan dalam proses **screening RAPBD** sebelum **disahkan**.