

# BEDAH DATA DJPK



**YOLO (You Only Live Once)**





# LATAR BELAKANG



# Latar Belakang

## Keputusan Menteri Dalam Negeri Nomor 050-5889 Tahun 2021

*“Klasifikasi, kodifikasi, dan nomenklatur perencanaan pembangunan dan keuangan daerah ditujukan untuk menyediakan informasi secara berjenjang melalui penggolongan, pemberian kode, dan daftar penamaan yang akan digunakan Pemerintah Daerah dalam perencanaan pembangunan dan keuangan daerah yang disusun secara sistematis sebagai pedoman dalam penyusunan dokumen perencanaan pembangunan daerah dan keuangan daerah”*

Terdapat **variasi nomenklatur** dari laporan keuangan daerah yang **belum sesuai dengan standar** yang telah ditetapkan.

### Numenklatur Tidak Sesuai

- Pengelolaan pendidikan khusus
- Rehabilitasi sedang/berat sarana, prasarana dan utilitas sekolah
- Penyelenggaraan Kejuaraan Olahraga

### Kesalahan Penulisan

- Penghargaan seniman dan budayawan\*\*\*\*)

### Numenklatur Kosong

- Nan

**Standarisasi nomenklatur dari laporan keuangan daerah perlu dilakukan!**

# Latar Belakang

*Pendidikan merupakan salah satu aspek penting pembangunan suatu negara dan menjadi fokus pemerintah, ditunjukkan dengan alokasi APBD yang besar.*

## ▶ UUD 1945 pasal 31 ayat (4)

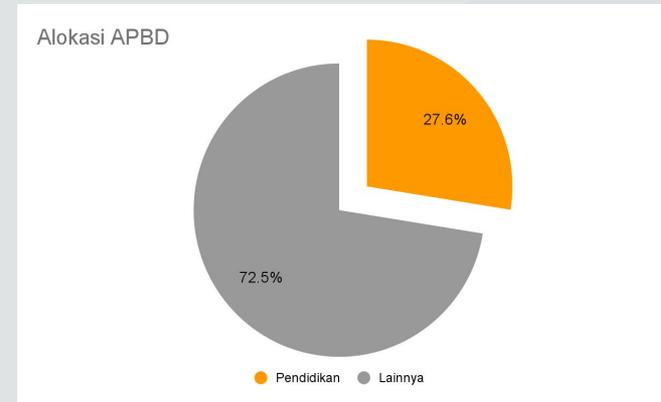
*“Negara memprioritaskan anggaran pendidikan **sekurang-kurangnya dua puluh persen** dari anggaran pendapatan dan belanja negara serta dari anggaran pendapatan dan belanja daerah untuk memenuhi kebutuhan penyelenggaraan pendidikan nasional”.*

## ▶ UU no 20 tahun 2003 tentang Sistem Pendidikan Nasional

### **Pasal 46**

(1) Pendanaan pendidikan menjadi **tanggung jawab bersama antara Pemerintah, Pemerintah Daerah**, dan masyarakat.

(2) Pemerintah dan Pemerintah Daerah bertanggung jawab menyediakan anggaran pendidikan sebagaimana diatur dalam Pasal 31 ayat (4) Undang- Undang Dasar Negara Republik Indonesia Tahun 1945



Hingga tanggal **21 Maret 2022**, pemerintah daerah telah membelanjakan anggaran tahun 2022 sebesar **27,55%** untuk fungsi pendidikan.

# Latar Belakang

*Adanya pengaruh besar APBD pada sektor pendidikan terhadap kualitas pendidikan, dan secara tidak langsung terhadap kesejahteraan rakyat*

▶ **Prasetyo, 2018**

*adanya pengaruh positif yang signifikan dari pengeluaran pemerintah di sektor pendidikan terhadap **angka melek huruf** dan **rata-rata lama sekolah**.*

▶ **Mongan, J. (2019) dan Sanniana Sidabutar, Elidawaty Purba, & Pawan Darasa Panjaitan. (2020).**

*pengeluaran pemerintah di sektor pendidikan berpengaruh positif terhadap nilai **Indeks Pembangunan Manusia**.*



*Dengan melihat bagaimana pengaruh pengeluaran pemerintah dalam bentuk APBD terhadap ketercapaian indikator pemerintah di tahun 2021, maka dapat diprediksi pula bagaimana ketercapaian indikator pendidikan di tahun 2022 berdasarkan APBD yang diajukan.*

# Tujuan & Output

- ▶ Melakukan **Standarisasi Nomenklatur** Subkegiatan dari Anggaran Pemerintah Daerah
- ▶ Membangun model untuk melihat **pengaruh alokasi APBD terhadap ketercapaian indikator pendidikan**, serta **mengestimasi nilai indikator pendidikan di tahun 2022** berdasarkan alokasi nilai APBD-nya
- ▶ Membangun **sistem/fitur usulan** untuk membantu DJPK dalam standarisasi nomenklatur laporan keuangan daerah.

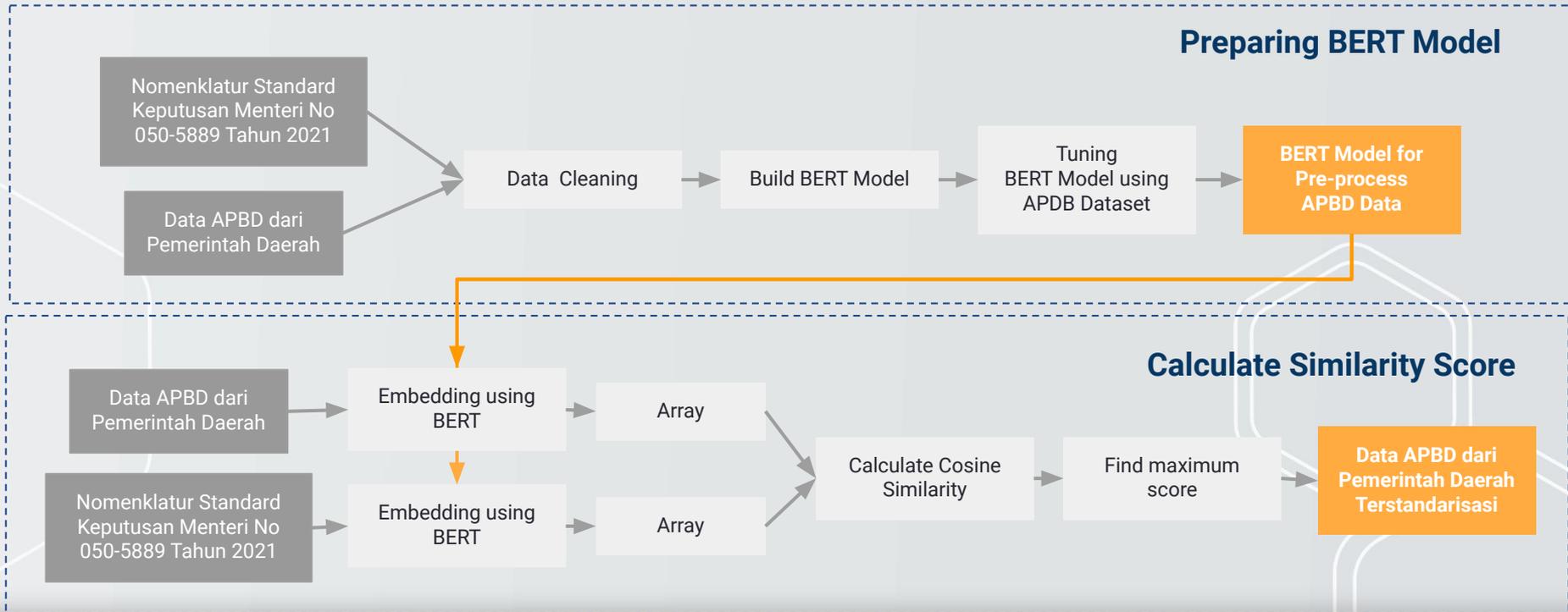


# **METODOLOGI**

## **Standarisasi & Analisis**



# Standarisasi Sub Kegiatan



# Standarisasi Sub Kegiatan

## SEMANTIC SEARCH

"Semantic search seeks to improve search accuracy by **understanding the content** of the search query. In contrast to traditional search engines, which only find documents based on **lexical matches**, **semantic search can also find synonyms.**"

### BERT

- **BERT** merupakan metode/mechanisme yang bertujuan untuk **mengkonversi input text menjadi numerik** agar dapat dilakukan pemodelan.
- **IndoBERT** merupakan versi indonesia dari model BERT. Model ini dilatih dengan menggunakan kata yang berasal dari: **Wikipedia bahasa Indonesia, artikel berita (Kompas, Tempo, and Liputan6)**, dan *Web Corpus Indonesia*.
- **Tuning** dilakukan untuk menjadikan pre-trained model BERT sesuai dengan kamus data yang akan digunakan, yaitu kata-kata berdasarkan data laporan keuangan APBD

### Cosine Similarity

- Pada saat pencarian, kueri dicocokkan dalam sebuah ruang vektor yang **sama dan terdekat** ditemukan dalam kamus.
- Isian ini seharusnya memiliki **kesamaan semantik/makna** yang tinggi dengan kueri yang dicari.
- Nilai kedekatan kedua kueri dihitung menggunakan rumus **cosine similarity**:

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

# Analisis | Regresi dan Interpretasi

## Regresi

- Regresi bertujuan untuk **mendapatkan pola hubungan dan mengukur seberapa kuat hubungan dari variabel bebas dengan variabel terikat**, serta untuk membangun model guna **memprediksi variabel terikat jika nilai variabel bebas diketahui**.
- Kami membandingkan beberapa metode regresi yaitu SVR, KNN, Decision Tree, Random Forest, ExtraTrees, XGBoost, dan Bagging Regressor.
- Metrics yang digunakan dalam pemilihan model terbaik adalah *Root Mean Square Logarithmic Error*

$$RMSLE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log(\hat{y}_i + 1) - \log(y_i + 1))^2}$$

## Interpretasi - SHAP Value

- Metode yang berdasarkan konsep kerjasama dalam game theory[2] yang biasa digunakan untuk **meningkatkan transparansi dan interpretabilitas** dari sebuah model machine learning.
- SHAP menunjukkan **kontribusi atau pengaruh** dari masing-masing feature data dalam memprediksi



**HASIL**



# Gambaran Data Standarisasi

Total Record APBD 2022

**5.242.575** record

Total Record dg Kode Sesuai

**18.231** [0,34%]

Total Record dg Kode Tidak Sesuai

**5.229.823** [99,75%]

Total Record Telah Sesuai Standar

**15.785** [0,30%]

Total Record Tidak Sesuai Standar

**2.446** [0,04%]

Proses standarisasi tidak dapat dilakukan menggunakan kode sub-kegiatan, melainkan menggunakan numenklatur sub-kegiatan.

# Hasil Semantic Search

Sebelum

Sesudah

kode_subkegiatan_ori	namasubkegiatan_clean
1.No.002.040000.070000	pengadaan mebel pendidikan nonformal kesetaraan
2.23.002.020000.010000	sosialisasi budaya baca dan literasi pada satuan pendidikan dasar dan pendidikan khusus serta masyarakat
1.No.005.020000.030000	pembinaan paud dan pendidikan nonformal yang diselenggarakan oleh masyarakat
1.No.002.040000.090000	pengadaan perlengkapan pendidikan nonformal kesetaraan
1.01.002.040000.180000	peningkatan kapasitas pengelolaan dana bos
1.01.002.040000.020000	pembangunan sarana prasarana dan utilitas sekolah nonformal kesetaraan
1.No.002.040000.050000	pemeliharaan rutin gedung ruang kelas ruang guru pendidikan nonformal kesetaraan

kode_subkegiatan_ref	nomenklatur_clean
1.01.020.204000.800000	pengadaan alat rumah tangga pendidikan non formal kesetaraan
1.01.060.201000.600000	peningkatan apresiasi siswa terhadap bahasa dan sastra daerah kewenangan kabupaten kota
1.01.050.202000.200000	pengendalian dan pengawasan perizinan paud dan pendidikan non formal yang diselenggarakan oleh masyarakat
1.01.020.204000.800000	pengadaan alat rumah tangga pendidikan non formal kesetaraan
1.01.020.201000.300000	peningkatan kapasitas pengelolaan dana bos sekolah dasar
1.01.020.204000.600000	pemeliharaan rutin sarana prasarana dan utilitas sekolah non formal kesetaraan
1.01.020.204000.300000	rehabilitasi sedang berat gedung ruang kelas ruang guru pendidikan non formal kesetaraan



**Model standarisasi yang dibangun tidak hanya dapat digunakan untuk nomenklatur sub-kegiatan, melainkan hierarki lainnya, seperti fungsi, program, kegiatan, dsb.**

# Hasil Standarisasi

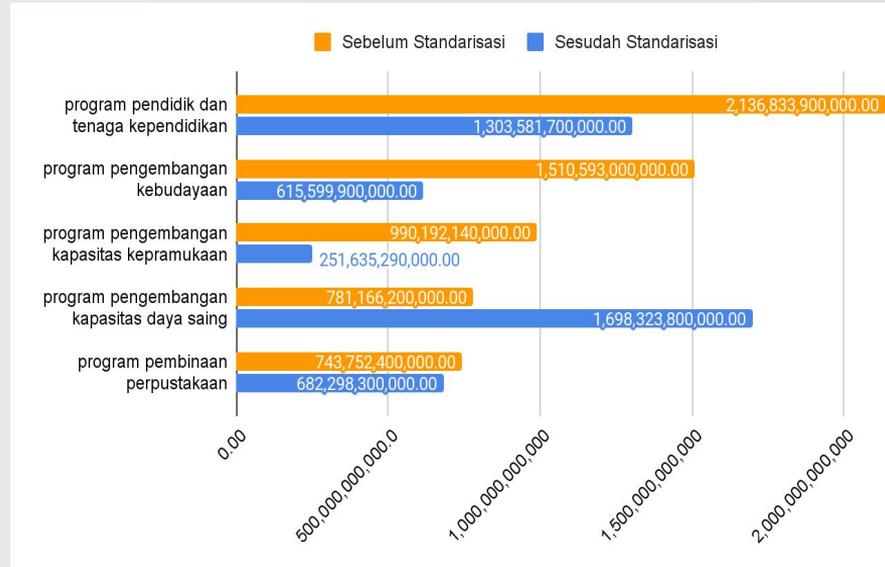
Kata Kunci:

***“Pembinaan Aparatur Pemerintah Desa”***

	nomenklatur	Cosine Similarity Score	kode_subkegiatan
0	Pembinaan Peningkatan Kapasitas Aparatur Pemerintah Desa	0.819340	2.13.040.201000.500000
1	Pemberdayaan Masyarakat di Kelurahan	0.761898	7.01.030.202000.300000
2	Pembinaan Sistem Informasi Pemerintahan Daerah di Bidang Pembangunan Daerah Pemerintah Kabupaten/Kota	0.749813	5.01.020.204000.300000
3	Pembinaan Perpustakaan Khusus Tingkat Kabupaten/Kota	0.743260	2.23.020.201000.500000
4	Penyusunan Kebijakan dan Strategi Daerah Pengelolaan Sampah Kabupaten/Kota	0.738740	2.11.110.201000.100000
5	Fasilitasi Manajemen Pemerintahan Desa	0.734924	2.13.040.201000.120000
6	Pelaksanaan Penilaian dan Evaluasi Kinerja Aparatur	0.731749	5.03.020.204000.200000
7	Pembinaan Akuntansi, Pelaporan dan Pertanggungjawaban Pemerintah Kabupaten/Kota	0.729958	5.02.020.203000.110000
8	Peningkatan Partisipasi Masyarakat dalam Forum Musyawarah Perencanaan Pembangunan di Desa	0.726774	7.01.030.201000.100000
9	Monitoring dan Evaluasi Pengelolaan Kawasan Strategis Pariwisata Kabupaten/Kota	0.724203	3.26.020.202000.500000

# Hasil Semantic Search

## Perbandingan Nilai Anggaran Sebelum dan Sesudah Standarisasi

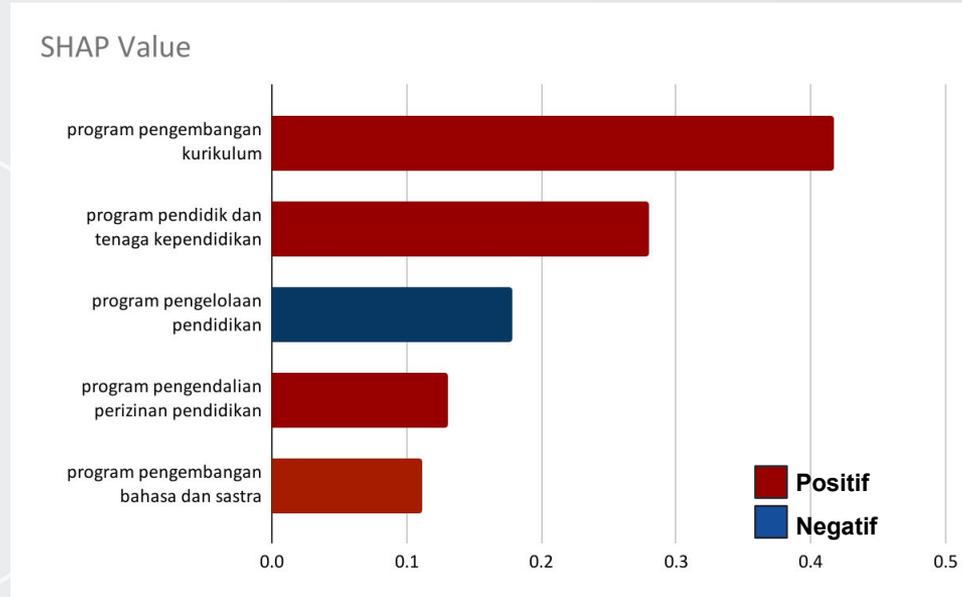


*“Tidak standarnya numenklatur dapat mengakibatkan kesalahan analisis dan pengambilan keputusan berkaitan dengan nilai APBD.”*

# Hasil Analisis [2]

## Model Regresi SVR

Nilai RMSLE : 0.19974



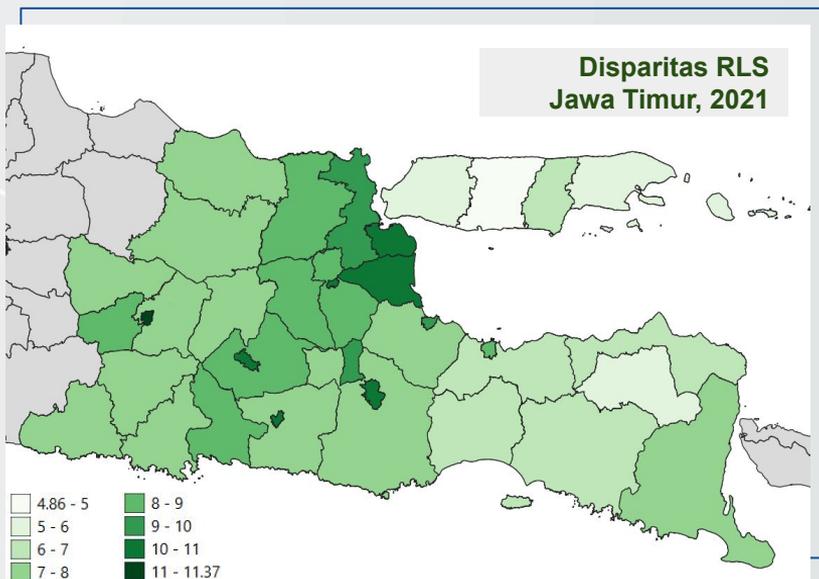
- Dua fitur yang berpengaruh signifikan adalah **program pengembangan kurikulum** dan **program pendidik dan tenaga kependidikan**, dengan **kontribusi positif** pada nilai RLS berdasarkan alokasi APBD-nya.
- **Semakin tinggi Alokasi APBD** untuk program pembangunan kurikulum, serta program pendidik dan tenaga pendidikan, diprediksi akan **semakin tinggi pula RLS** di daerah tersebut.
- **Kualitas kurikulum dan tenaga pendidik** merupakan dua kunci keberhasilan pendidikan dalam mencetak generasi-generasi terdidik. Semakin tinggi alokasi APBN mencerminkan prioritas daerah dalam mengembangkan sektor pendidikan.



**Optimalkan anggaran pada program dengan pengaruh positif terhadap RLS**

# Cakupan Analisis

## Cakupan Analisis: Provinsi Jawa Timur



RLS 2021

7,88



Meskipun Jawa Timur memiliki banyak sekolah bagus, tetapi merupakan provinsi dengan rata-rata lama sekolah **terendah ke-7**, dibawah rata-rata nasional sebesar 8,54

Range RLS

6,51



Masih adanya ketimpangan rata-rata lama sekolah kabupaten/kota di Jawa Timur, dengan **selisih nilai RLS tertinggi dan terendah terjauh ke-2** setelah Provinsi Papua. RLS tertinggi berada di **Kota Madiun** dan terendah **Kabupaten Sampang**.

APBD 2021



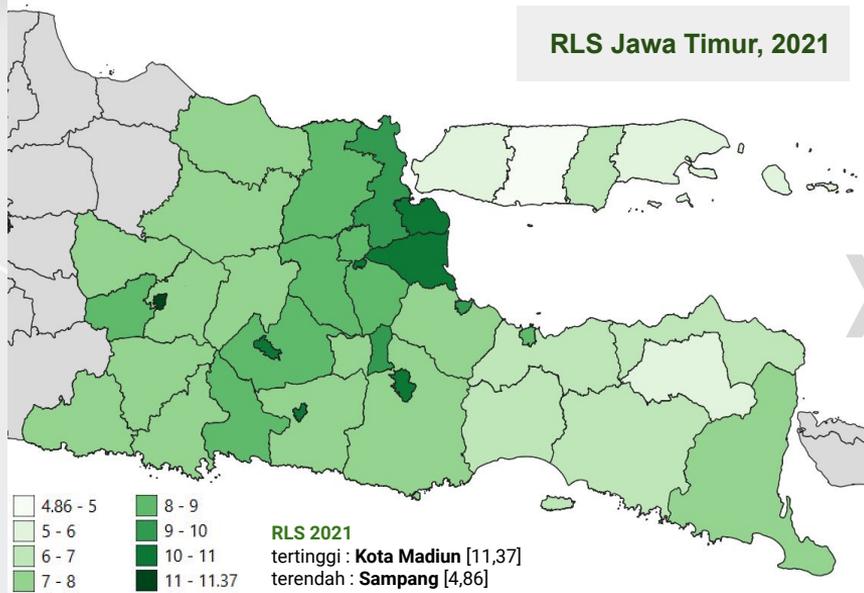
Total APBD **tertinggi no-1** di Indonesia  
Jumlah APBD yang terealisasi **tertinggi no-1** di Indonesia,



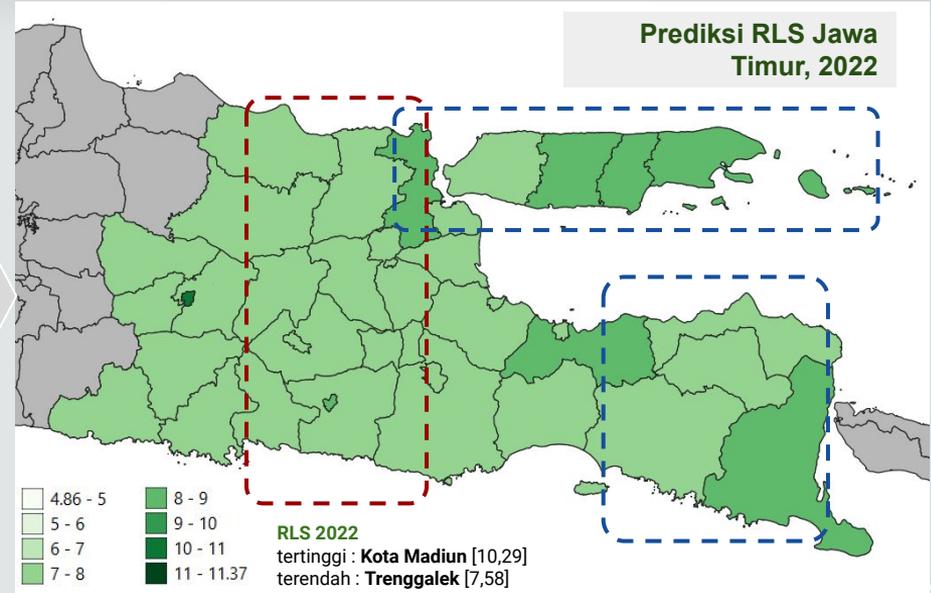
Adanya relasi terbalik antara APBD di Provinsi Jawa Timur dengan outcome yang diharapkan (rata-rata lama sekolah) menarik untuk diteliti lebih lanjut :)

# Prediksi Indikator Pendidikan Tahun 2022

RLS Jawa Timur, 2021



Prediksi RLS Jawa Timur, 2022



➤ **18 Kabupaten/Kota** diprediksi mengalami peningkatan nilai Rata-rata Lama Sekolah, yaitu area sebelah timur (**Situbondo, Bondowoso, Jember, dan Banyuwangi**), dan Pulau Madura (**Sumenep, Pamekasan, Sampang dan Bangkalan**)

➤ **20 Kabupaten/Kota** diprediksi mengalami penurunan nilai Rata-rata Lama Sekolah, diantaranya **Malang, Lamongan, Mojokerto, Jombang, Kediri, dan Tulungagung**

# Tinjauan Khusus: Perbandingan Target dan Hasil Prediksi

## Kabupaten Malang



RLS 2021: 7,43

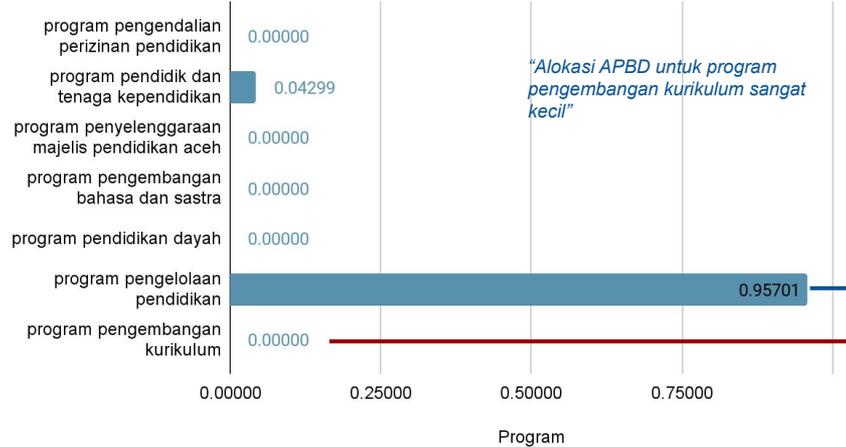


Target RLS2022 : 14,6

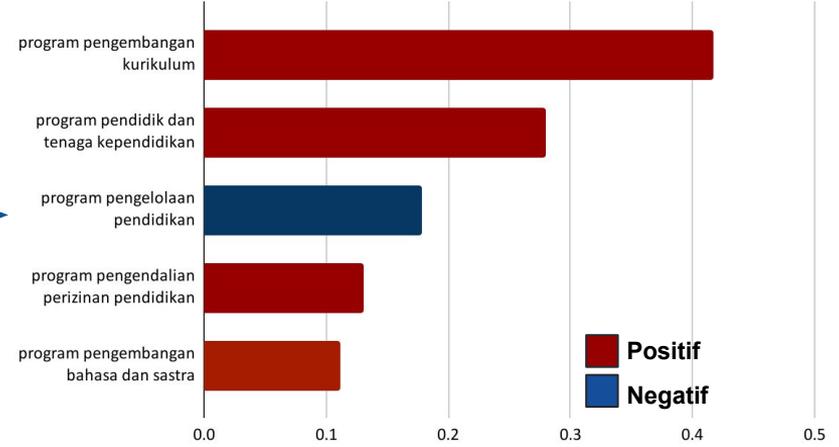


Prediksi RLS 2022 : 7,426

### Rasio Anggaran per Program



### SHAP Value

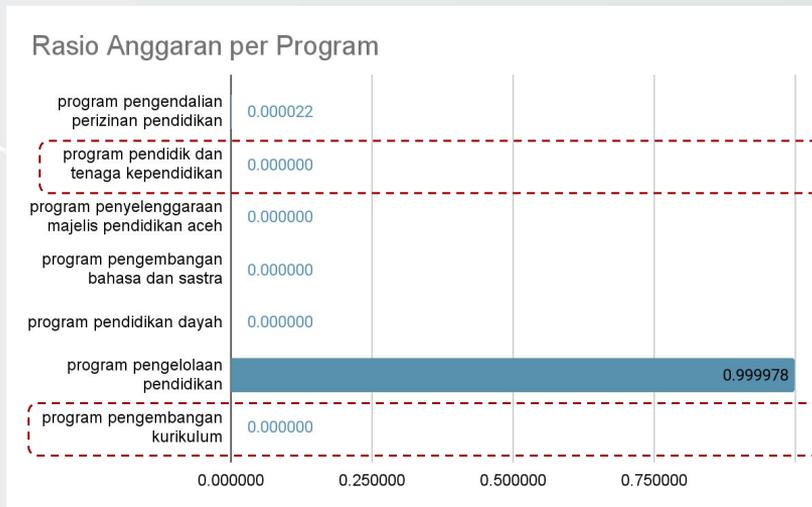


Me-optimalkan alokasi APBD **program pengembangan kurikulum** dan/atau **program pendidik dan tenaga kependidikan** yang memiliki pengaruh positif pada nilai RLS

# Tinjauan Khusus: Perbandingan Target dan Hasil Prediksi [2]

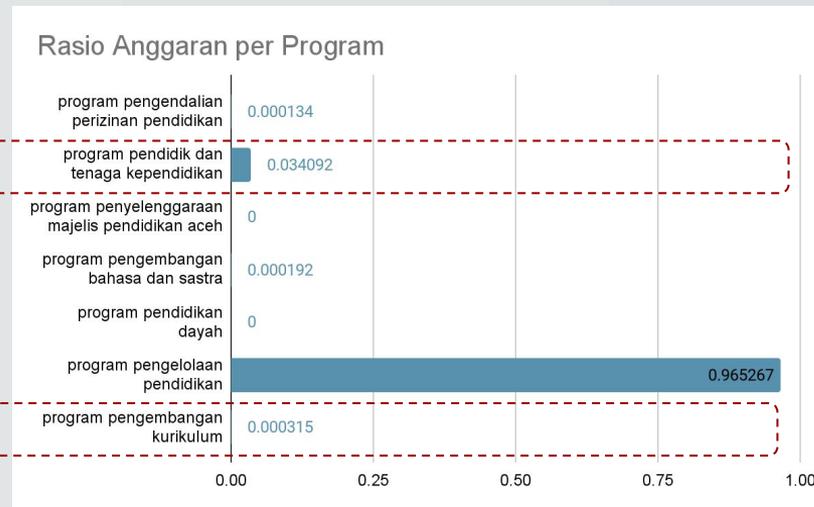
## Kabupaten Sidoarjo

RLS 2021: 10.72 ↓ Prediksi RLS 2022 : 7.659



## Kabupaten Sampang

RLS 2021: 4.86 ↑ Prediksi RLS 2022 : 8.079



Me-optimalkan alokasi APBD program pengembangan kurikulum dan/atau program pendidik dan tenaga kependidikan yang memiliki pengaruh positif pada nilai RLS

# Rekomendasi Kebijakan

- **Mengoptimalkan alokasi anggaran pada program-program yang berpengaruh positif** terhadap nilai indikator pendidikan sekolah seperti program pengelolaan pendidikan.
- Membuat **batas/limit maksimal pengeluaran APBD** pada program-program yang berpengaruh negatif terhadap nilai indikator pendidikan, sehingga daerah dapat memfokuskan anggarannya pada program-program lain yang lebih bermanfaat. Batas/limit dibuat dalam format persentase terhadap total anggaran pada kabupaten/kota terkait.
- Mensosialisasikan dan **menganjurkan Standict App** untuk digunakan di pemerintah daerah kabupaten maupun provinsi dalam penyusunan APBD dan **mengimplementasikan Standict Model** di Sistem Informasi Keuangan Daerah (SIKD).

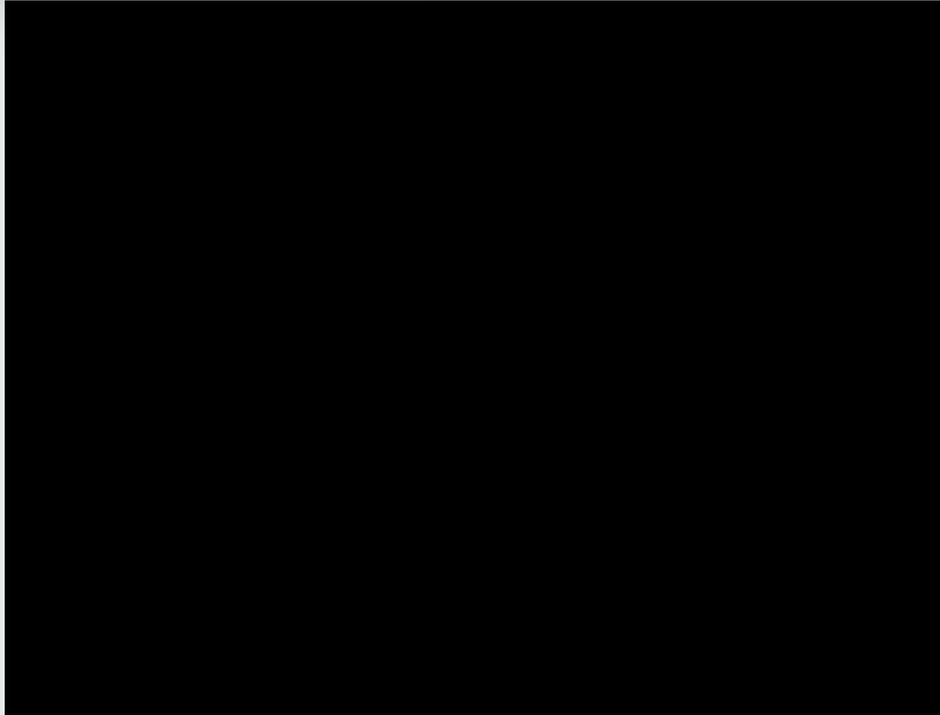


# **SISTEM USULAN**



# Sistem Usulan - Demo

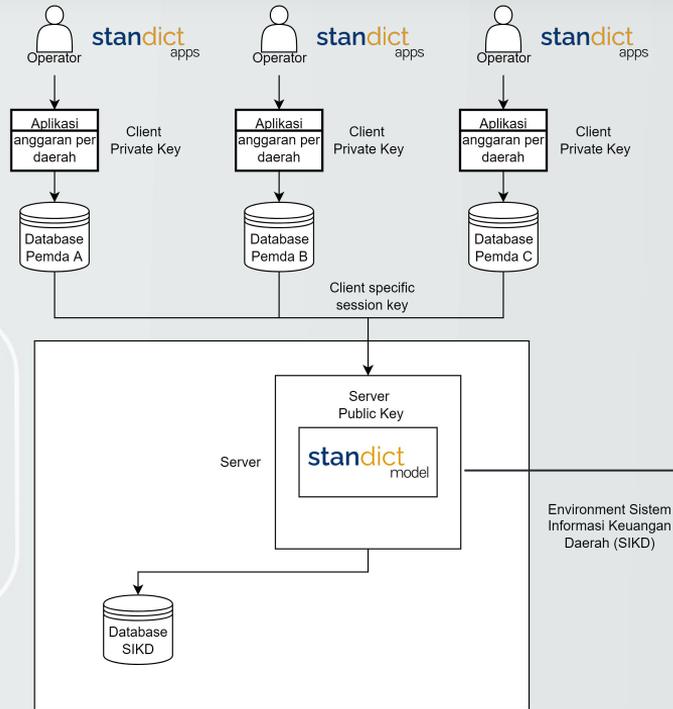
standict  
apps



standict  
model

[https://drive.google.com/file/d/1jDQp-hNDA8Epg4XW\\_sLSh8tDCHOqIYqY/view?usp=sharing](https://drive.google.com/file/d/1jDQp-hNDA8Epg4XW_sLSh8tDCHOqIYqY/view?usp=sharing)

# Arsitektur Sistem Usulan



Standict Apps membantu Operator atau penyusun APBD di daerah dalam menentukan nomenklatur sub kegiatan yang baku

Standict Model membantu DJPK untuk memvalidasi dan memperbaiki subkegiatan dari data APBD yang telah disubmit oleh daerah agar sesuai dengan standar sebelum masuk ke Database SIKD



Memudahkan pemerintah daerah dalam menyusun rancangan anggaran yang sesuai dengan nomenklatur standar sesuai peraturan yang berlaku.

# Future Development

- Penelitian ini hanya menggunakan satu contoh outcome, yaitu RLS (Rata-rata lama sekolah) di Provinsi Jawa Timur. Selanjutnya, pemodelan dapat dilakukan **untuk masing-masing outcome**, sehingga dapat mencakup seluruh indikator yang ingin dicapai dan memperkirakan bagaimana alokasi optimum agar output tersebut tercapai.
- Menambahkan **sistem auto-fill atau sistem rekomendasi** yang bekerja ketika penyusun rencana anggaran mengetik usulan nomenklatur pada sistem.

# Closing Statements



*"Anggaran yang dikelola dengan baik tidak hanya mencerminkan kualitas ekonomi yang baik, tapi mencerminkan martabat suatu bangsa yang baik."*

*"Sri Mulyani"*

# Terima Kasih!

#DataMencerdaskanBangsa  
#MencatatPertanianIndonesia

#MenjagaKeuanganNegara