The background is a dark navy blue. In the top-left corner, there are two overlapping geometric shapes: a blue parallelogram and a light green parallelogram. In the bottom-left corner, there is a circular inset showing a detailed, grayscale image of a circuit board. In the top-right corner, there is a faint, grayscale image of a circuit board with many small components.

# AwanTunai Credit Analytics Mission



# Overview

AwanTunai merupakan marketplace yang menyediakan fasilitas cicilan dengan 15 menit cair, dengan menggunakan jaringan merchant offline, dengan mitra:

- Perusahaan (dengan target karyawan maupun konsumen), Bersifat Custom
- Mitra toko HP, Limit 5jt
- Mitra Lain (Toko Susu, Bengkel, Apotik, Toko Bangunan) limit 750rb
- Pinjaman pendidikan

# Understanding the problems

- 01 Meningkatkan penetrasi smartphone, karena saat ini smartphone digunakan sebagai alat mencari nafkah.
- 02 Penduduk Bankable di Indonesia masih 37% , dan hanya tumbuh 8-10% per tahun
- 03 Biaya perbankan konvensional sangat tinggi, khususnya di sektor kredit, biasa survey, underwriting, kostumisasi kebutuhan, dll



# Solusi



# What We Have

Masalah yang sangat jelas dalam bahasa bisnis

Tindakan yang akan dilakukan jika output telah selesai

Adaaya spesifikasi dan personalisasi dari setiap tindakan

Pemahaman tentang pengkonsumsian machine learning sampai tingkat eksekutor langsung





Persona 01

# Misi

1. Analisis Credit
2. Deteksi Aplikasi Fraud (Merchant, maupun dari Customer)
3. Efisiensi Penagihan
4. Optimasi Limit Peminjaman
5. Verifikasi Identitas
6. Prediksi Income
7. Prediksi Nama
8. Segmentasi Customer
9. Rekomendasi Merchant
10. dll



# What we need

---

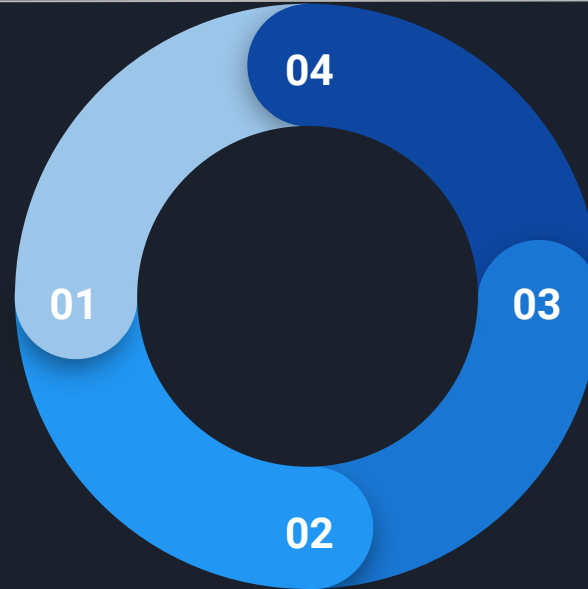
Data and Engine

Support and  
Feedback

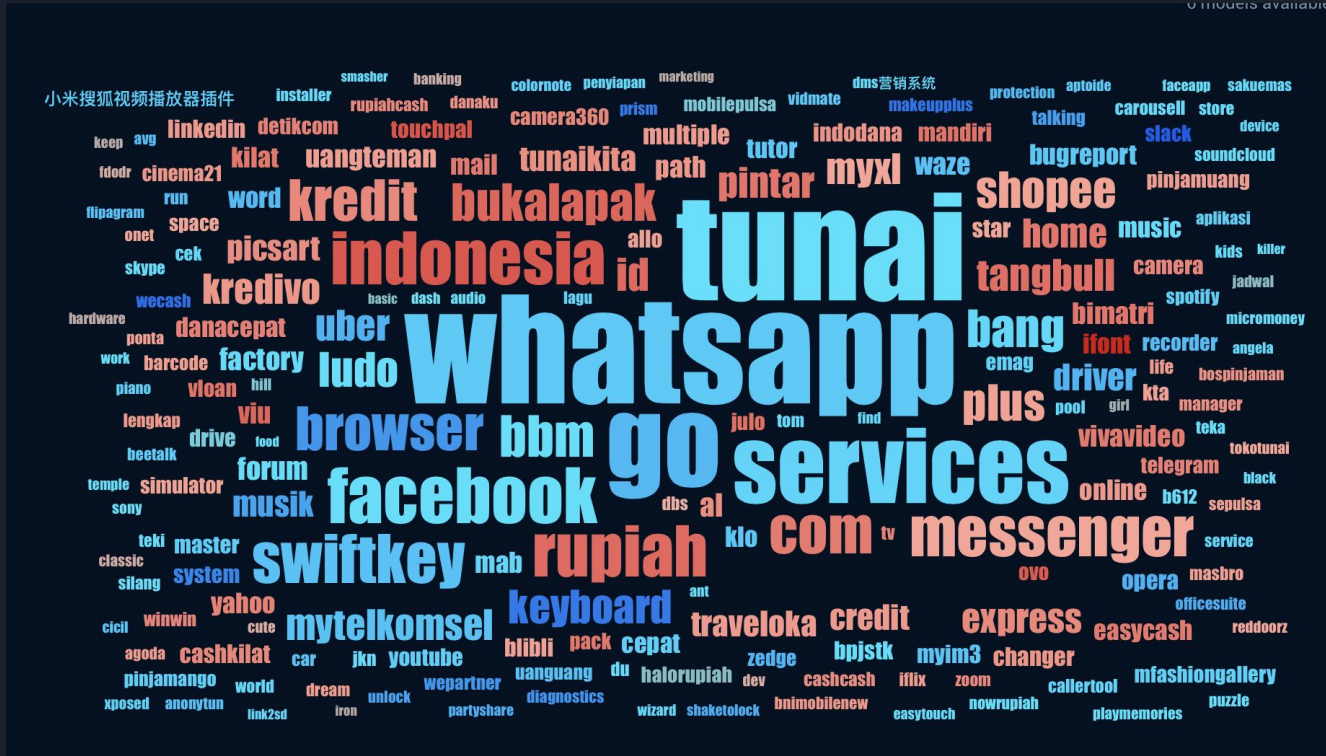
Talent

Execution  
Technology

---



# Feature Engineering







# Feature Engineering

## NUMERIC

Imputasi (Median,  
Mean)

Monotomic  
Transformation (Square,  
Log, Ridit, Logit, Probit,  
Box-Cox, etc)

Diferensi

Rasio

## KATEGORIK

One-Hot Encoding

Label Encoding

Count Encoding

Custom Binary Encoding

Buhlmann Credibility  
Estimates

BackwardDifferenceEnc  
oder

## Text

N-Grams

TF-IDF

Uysal-Gunal

Cosine Similarity

Tokenization

Grouping Tokenization



Long-Lat

Vicinity Distance

Radius Fill

Redflag/Blueflag

To Categorical

To Numeric

To Text

Timestamp

Year, Month, week

Day,, Hour, Minutes,  
Second

Duration

Intensity

Interval

Mix

Image

Gaussian Blurr

Grayscale

CNN

Size

Colour

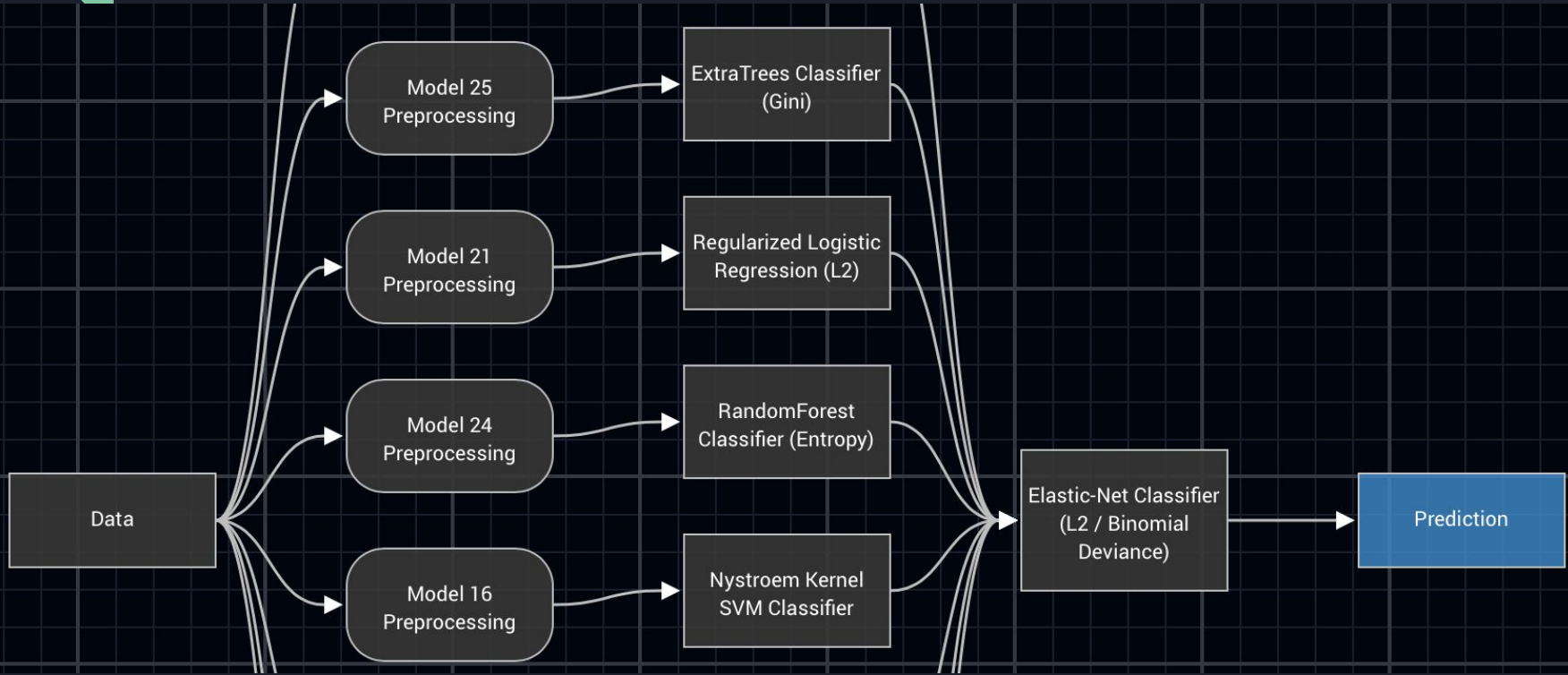
Blurriness

etc

# Showcase

| Feature Name   | Index | Importance ▾           | Var Type | Unique | Missing | Mean | Std Dev | Mediar |
|--|-------|------------------------|----------|--------|---------|------|---------|--------|
| <input type="checkbox"/> no_credit_approvals         | 31    | <div><div></div></div> | Numeric  | 7      | 55      | 0.23 | 0.85    | 0      |
| <input type="checkbox"/> government                  | 8     | <div><div></div></div> | Numeric  | 10     | 0       | 0.90 | 1.40    | 0      |
| <input type="checkbox"/> monthly_avg_...ce_3am_calls | 32    | <div><div></div></div> | Numeric  | 62     | 79      | 1.71 | 2.95    | 1      |
| <input type="checkbox"/> negative_sms_count_a2m      | 27    | <div><div></div></div> | Numeric  | 11     | 56      | 0.29 | 1.30    | 0      |
| <input type="checkbox"/> food                        | 7     | <div><div></div></div> | Numeric  | 55     | 0       | 6.02 | 13.86   | 0      |
| <input type="checkbox"/> incoming                    | 2     | <div><div></div></div> | Numeric  | 404    | 0       | 457  | 822     | 196    |
| <input type="checkbox"/> transport                   | 14    | <div><div></div></div> | Numeric  | 56     | 0       | 7.19 | 16.92   | 1      |

# ML - Architecture





## Model Overview

 MODEL FILE SIZE

**10.255 MB**

 PREDICTION TIME

**54.0662s** 

Time to score 1,000 rows

 SAMPLE SIZE

**589 rows**

Training 471 rows  
Test 118 rows

PARTITION

WALL CLOCK TIME

CV #1

13.2 s

CV #3

13.2 s

CV #2

13.2 s

CV #5

13.2 s

CV #4

13.2 s

# Example

Journal of Machine Learning Technologies  
ISSN: 2229-3981 & ISSN: 2229-399X, Volume 2, Issue 1, 2011, pp-37-63  
Available online at <http://www.bioinfo.in/contents.php?id=51>



## EVALUATION: FROM PRECISION, RECALL AND F-MEASURE TO ROC, INFORMEDNESS, MARKEDNESS & CORRELATION

○ Maximizes F1 Score    - NOT(Good)    + Good

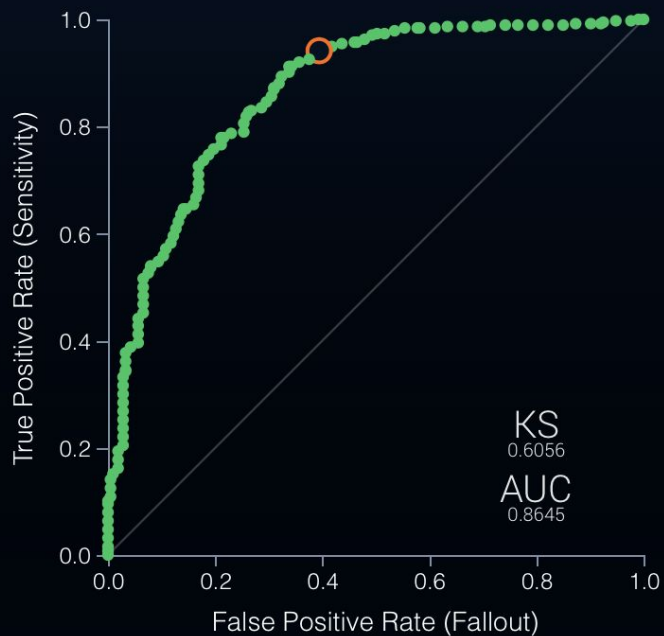
| F1 Score | True Positive Rate<br>(Sensitivity) | False Positive Rate<br>(Fallout) | True Negative Rate<br>(Specificity) | Positive Predictive<br>Value (Precision) | Negative Predictive<br>Value | Accuracy | Matthews Correlation<br>Coefficient |
|----------|-------------------------------------|----------------------------------|-------------------------------------|--|------------------------------|----------|-------------------------------------|
| 0.8698   | 0.9415                              | 0.3944                           | 0.6056                              | 0.8082                                   | 0.8543                       | 0.82     | 0.6021                              |
| Actual   |                                     | Predicted                        |                                     |  |                              |          |                                     |
|          |                                     | -                                |                                     | +  |                              |          |                                     |
|          |                                     | -                                |                                     | 213                                      |                              |          |                                     |
|          |                                     | +                                |                                     | 376                                      |                              |          |                                     |
|          |                                     | 151                              |                                     | 438                                      |                              | 589      |                                     |

# Chart

## ROC Curve

Data Source:

Cross Validation ▼



## Prediction Distribution

Threshold (0-1):

0.4664

Density ▼

