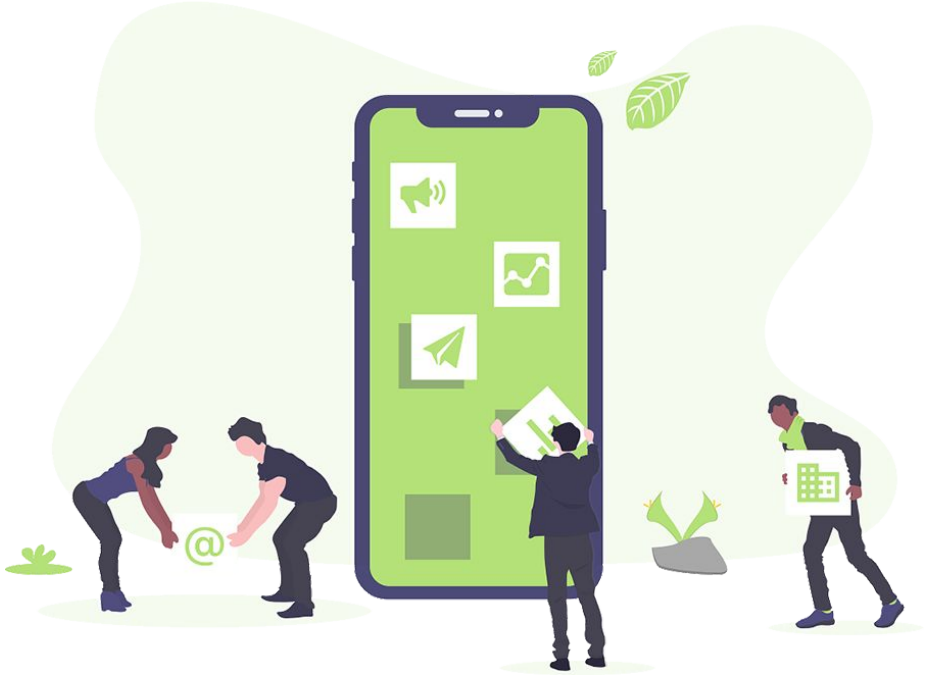
Handwritten Name Recognition using Tensorflow

Adi Kurniawan - 152236035100-987
Dikky Kurniawan - 152236035100-51
Elyadityo Santoso - 152236035101-567
Novri Kurniawan - 152236035100-1068
Robby Alfardo Irfan - 152236035100-380



Machine Learning with TensorFlow Training
Professional Academy Digital Talent Scholarship 2022

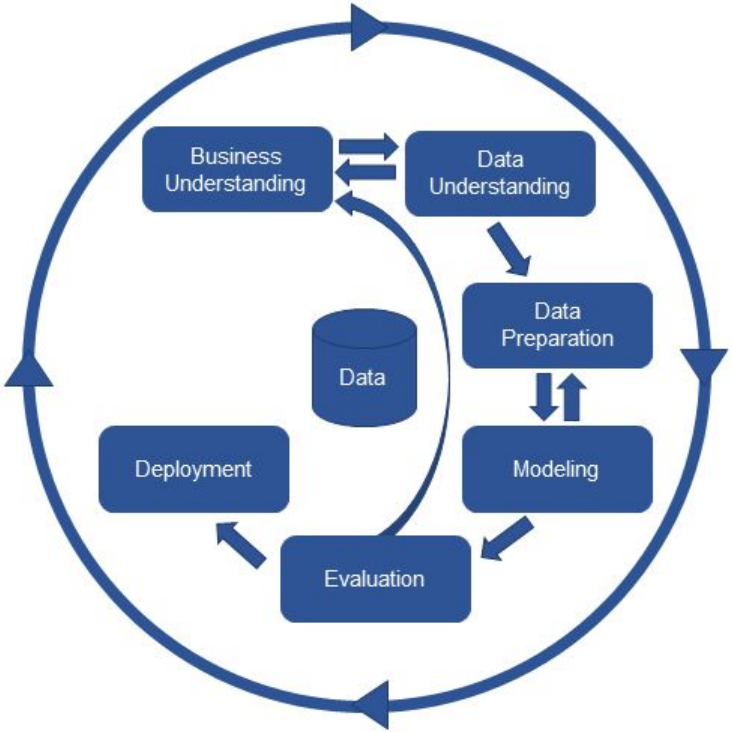
1. Project Framework

techready PRO Professional Academy digitalent KOMINFO

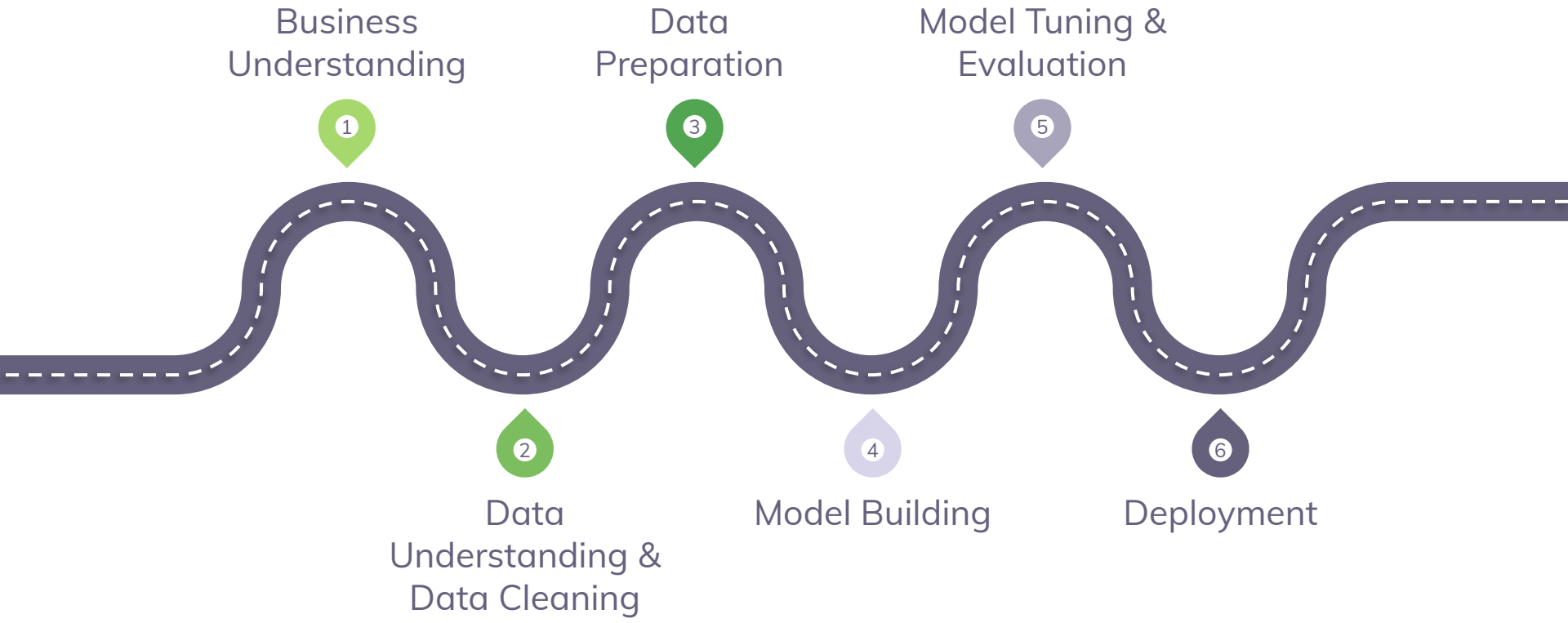
Framework

CRISP – DM

CRoss **I**ndustry **S**tandard **P**rocessing for **D**ata **M**ining



Roadmap


Business Understanding
1
Data Understanding & Data Cleaning
2
Data Preparation
3
Model Building
4
Model Tuning & Evaluation
5
Deployment
6

Project Schedule

Task	Juli								
	9-10	11-12	13-14	15-16	17-18	19-20	21-22	23-24	25-27
Perkenalan masing-masing anggota									
Membentuk dan menyelesaikan proposal									
Mencari dan mempelajari paper terkait handwriting recognition									
Explore Handwriting Dataset									
Cleaning Handwriting Dataset									
Melakukan Exploratory Data Analysis									
Melakukan Modelling TensorFlow									
Membandingkan hasil model masing-masing anggota									
Mempersiapkan PPT hasil Modelling									
Presentasi hasil model dan pembuatan video									

2. Business Understanding

techready PRO Professional Academy digitalent KOMINFO

Business Understanding

techready PRO Professional Academy digitalent KOMINFO

Business Problem

Top Companies

Google amazon

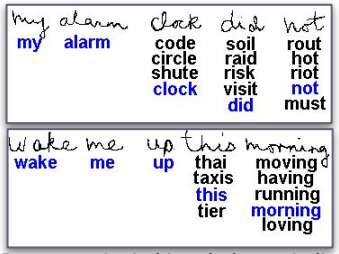
Data

Data merupakan **fakta yang tidak terorganisir** dan **butuh untuk diproses.**

Contoh data yaitu, **angka, teks, suara, gambar, animasi, dan video.**

Information

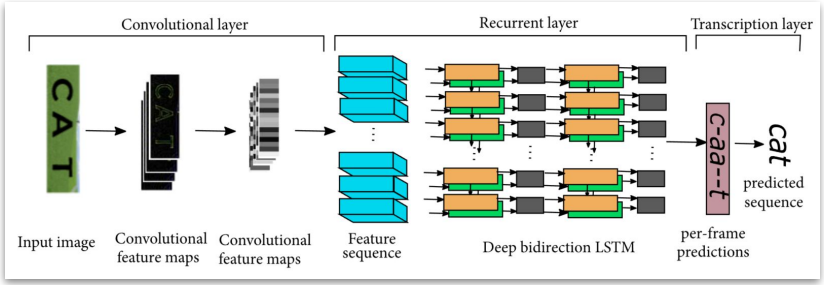
Ketika data **terproses, terorganisir, dan terstruktur,** itu akan membuat lebih **berguna.**



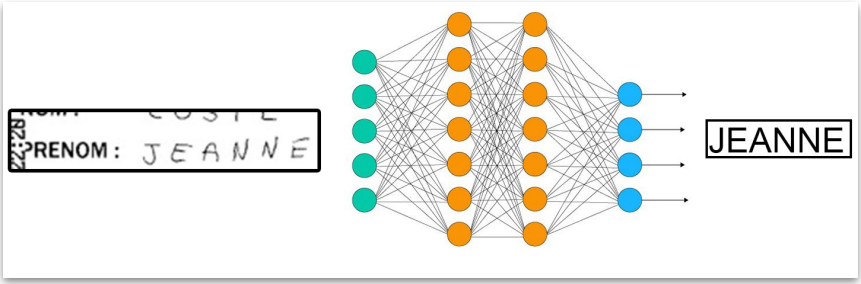
Data yang ingin kita ubah menjadi **informasi** adalah **tulisan tangan manusia.**

Business Solution

Menggunakan Skema Convolutional Recurrent Neural Network (CRNN)



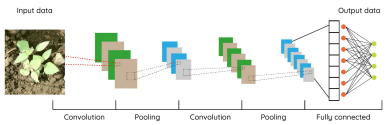
Contoh case Handwritten Names Recognition untuk menginterpretasi tulisan tangan pada gambar.



Business Objectives

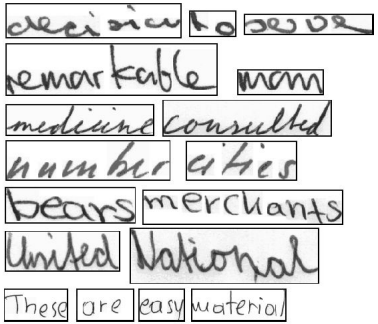
Dengan mengembangkan model ini ada beberapa hal yang ingin tim kami capai:

Mengenali Pengaplikasian Tensorflow



TensorFlow

Mengeksplorasi Dataset



Mempelajari Cara Berpikir Terstruktur

Model Menjadi Dasar Pengembangan Selanjutnya

3. Data Understanding

techready PRO Professional Academy digitalent KOMINFO

Dataset

written_name_train_v2				
Columns	**Unique values**	**Missing Values**	**%Missing Values**	**Example**
FILENAME	330.961	0	0,00%	TRAIN_00001.jpg
IDENTITY	100.540	565	0,56%	BALTHAZAR

written_name_validation_v2				
Columns	**Unique values**	**Missing Values**	**%Missing Values**	**Example**
FILENAME	41.370	0	0,00%	VALIDATION_0001.jpg
IDENTITY	20.228	78	0,39%	BILEL

written_name_test_v2				
Columns	**Unique values**	**Missing Values**	**%Missing Values**	**Example**
FILENAME	41.370	0	0,00%	TEST_0001.jpg
IDENTITY	20.280	70	0,35%	KEVIN

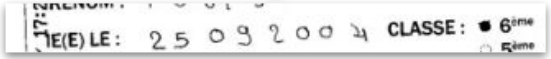
Exploratory Data Analysis

Contoh Image dengan jumlah karakter 1:

-

A

D



A

-

X

Contoh Image dengan jumlah karakter 2:

LE

LY

BE

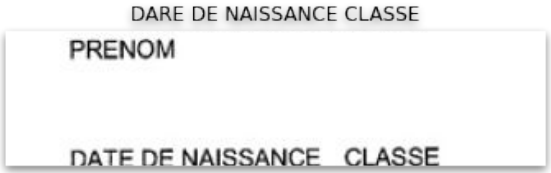
XU

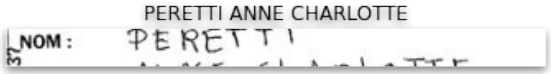
LE

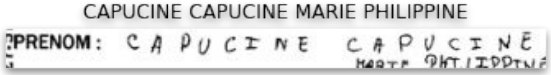
KY

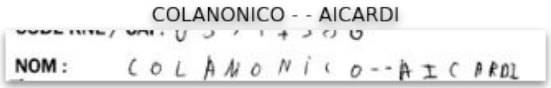
Exploratory Data Analysis

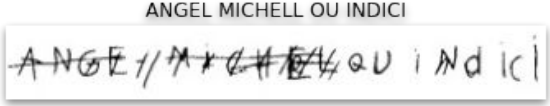
Contoh Image dengan jumlah karakter lebih dari 21:











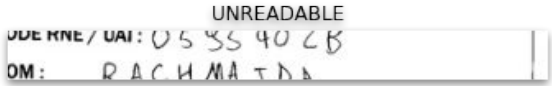
Contoh Image dengan Identity “-” dan “--”:

Exploratory Data Analysis

Contoh Image dengan Identity “UNREADABLE”:

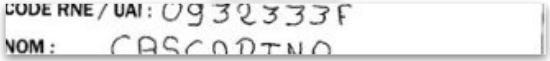
UNREADABLE

UNREADABLE

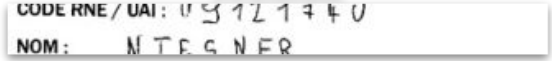


UNREADABLE

UNREADABLE



UNREADABLE



UNREADABLE

4. Data Preparation

- Data Cleaning
- Data Preparation

Data Cleaning

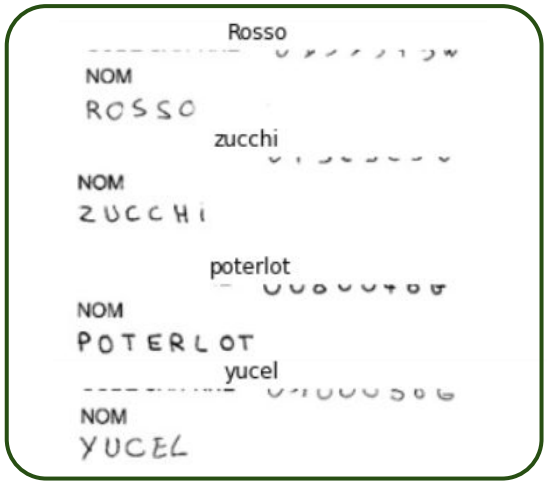
Missing Label

Jumlah:
565 pada training, 78 pada validation, dan 70 pada testing.

1. pelabelan manual -> tidak memungkinkan
2. Jumlahnya tidak signifikan dibanding total sampel
3. Variasi image pada variabel NA masih terwakili oleh sampel lain

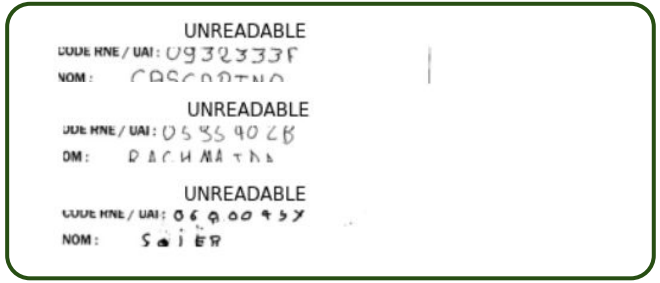
Treatment:
Dikeluarkan dari sampel

Non Capital Label



Treatment:
Kesalahan Label, sehingga diubah menjadi capital

Unreadable Label



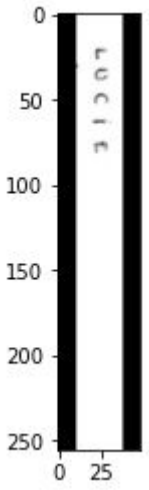
1. Terdapat inkonsistensi data yang dilabeli unreadable.
2. Telah dicoba membuat model untuk predict readability, tapi kurang optimal

Treatment:
Dikeluarkan dari training set. Untuk test tetap dimasukkan

Data Preparation

Resize Gambar

1. Gambar perlu dipastikan memiliki dimensi yang sama
2. Resize gambar ke ukuran 256x64 (WxH)
3. Rotasi 90 derajat searah jarum jam agar lebih mudah diekstraksi fiturnya



Padding Label

pengecekan terhadap karakter yang digunakan pada label yang telah di ubah ke huruf kapital

Characters: [' ', "'", '-', 'A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J', 'K', 'L', 'M', 'N', 'O', 'P', 'Q', 'R', 'S', 'T', 'U', 'V', 'W', 'X', 'Y', 'Z']

Set max input label = 16
Label > 16 = tidak digunakan di training
Label < 16 = dipadding degan spesial token untuk melengkapi 16

Data Pipeline

Mengoptimalisasikan proses operasi terutama pada dataset gambar yang membutuhkan *resource* yang besar

1. tf.data.Dataset.from_tensor_slices
2. tf.io.read_file(image_path)
3. tf.image.resize_with_pad
4. tf.keras.layers.StringLookup
5. dataset.batch(batch_size).cache()
   .prefetch(tf.data.AUTOTUNE)

5. Model Building

techready PRO Professional Academy digitalent KOMINFO

Model Building

Model Building:

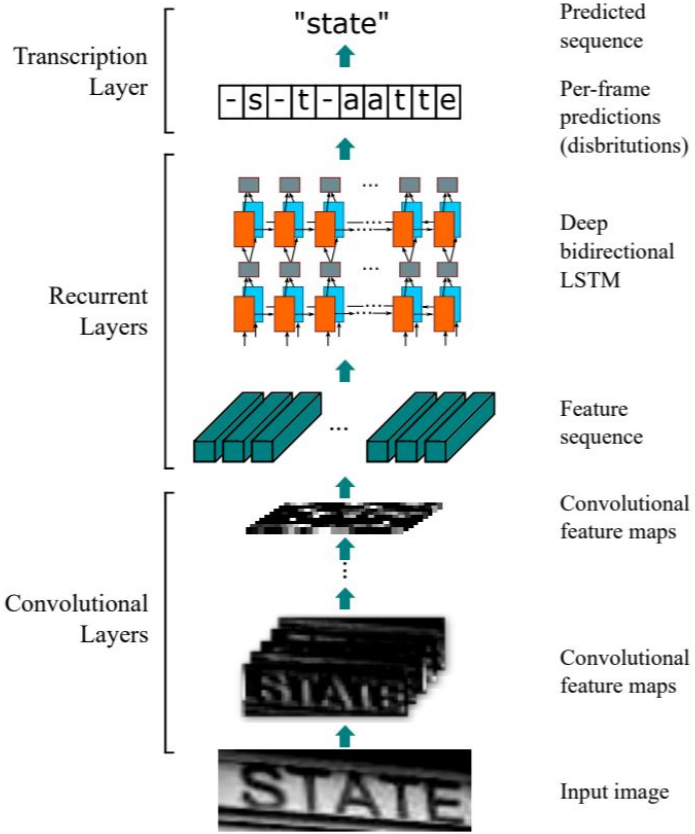
Convolutional Neural Network:
Mengektraksi fitur dari gambar untuk tiap karakter

Recurrent Neural Network:
Menentukan sequential dari karakter untuk membentuk kata

CTC Loss Function:
Menentukan probabilistik dari kata yang diprediksi dengan label

Transcription:

CTC Decoding:
greedy / best path decoding



Sumber gambar: http://www.sxlwork.top/

Model Building

Model Building:

BASE MODEL

Convolutional Neural Network:
2 Layer konvolusi 3x3
2 Layer Max Pooling

Recurrent Neural Network:
2 Layer bidirectional LSTM
Output berupa kelas character vocabulary

CTC Loss Function

Model ini mengadopsi dari keras example:
https://keras.io/examples/

```
Model: "handwriting_recognizion"
__________________________________________________________________________________________________
Layer (type)                    Output Shape         Param #     Connected to
==================================================================================================
image (InputLayer)              [(None, 64, 256, 1)] 0
__________________________________________________________________________________________________
Conv1 (Conv2D)                  (None, 64, 256, 32)  320         image[0][0]
__________________________________________________________________________________________________
pool1 (MaxPooling2D)            (None, 32, 128, 32)  0           Conv1[0][0]
__________________________________________________________________________________________________
Conv2 (Conv2D)                  (None, 32, 128, 64)  18496       pool1[0][0]
__________________________________________________________________________________________________
pool2 (MaxPooling2D)            (None, 16, 64, 64)   0           Conv2[0][0]
__________________________________________________________________________________________________
reshape (Reshape)               (None, 64, 1024)     0           pool2[0][0]
__________________________________________________________________________________________________
dense1 (Dense)                  (None, 64, 64)       65600       reshape[0][0]
__________________________________________________________________________________________________
dropout_2 (Dropout)             (None, 64, 64)       0           dense1[0][0]
__________________________________________________________________________________________________
bidirectional_4 (Bidirectional) (None, 64, 256)      197632      dropout_2[0][0]
__________________________________________________________________________________________________
bidirectional_5 (Bidirectional) (None, 64, 128)      164352      bidirectional_4[0][0]
__________________________________________________________________________________________________
label (InputLayer)              [(None, None)]       0
__________________________________________________________________________________________________
dense2 (Dense)                  (None, 64, 31)       3999        bidirectional_5[0][0]
__________________________________________________________________________________________________
ctc_loss (CTCLayer)             (None, 64, 31)       0           label[0][0]
                                                                 dense2[0][0]
==================================================================================================
Total params: 450,399
Trainable params: 450,399
Non-trainable params: 0
__________________________________________________________________________________________________
```

6. Model Tuning & Evaluation

techready PRO Professional Academy digitalent KOMINFO

Single Evaluation Metrics

Traditional Metrics
Akurasi, presisi, recall

→ **Tidak Cocok DIgunakan untuk OCR/Text Recognition**

Rekomendasi:
Menggunakan Character Error Rate (CER) atau Word Error Rate (WER)

→ CER -> jumlah karakter yang yg harus diubah untuk membentuk label yang benar dibagi dengan jumlah karakter
CER -> cocok untuk urutan karakter berjumlah sedikit
WER -> cocok untuk transkripsi paragraf atau halaman

Dataset:
1-2 kata, jumlah karakter <20
Lebih cocok untuk dideteksi berdasarkan karakter level

→ **Menggunakan CER sebagai pedoman dalam model tuning dan mengevaluasi model**

Bias Optimal Error

Jumlah Error minimal yg dapat dicapai

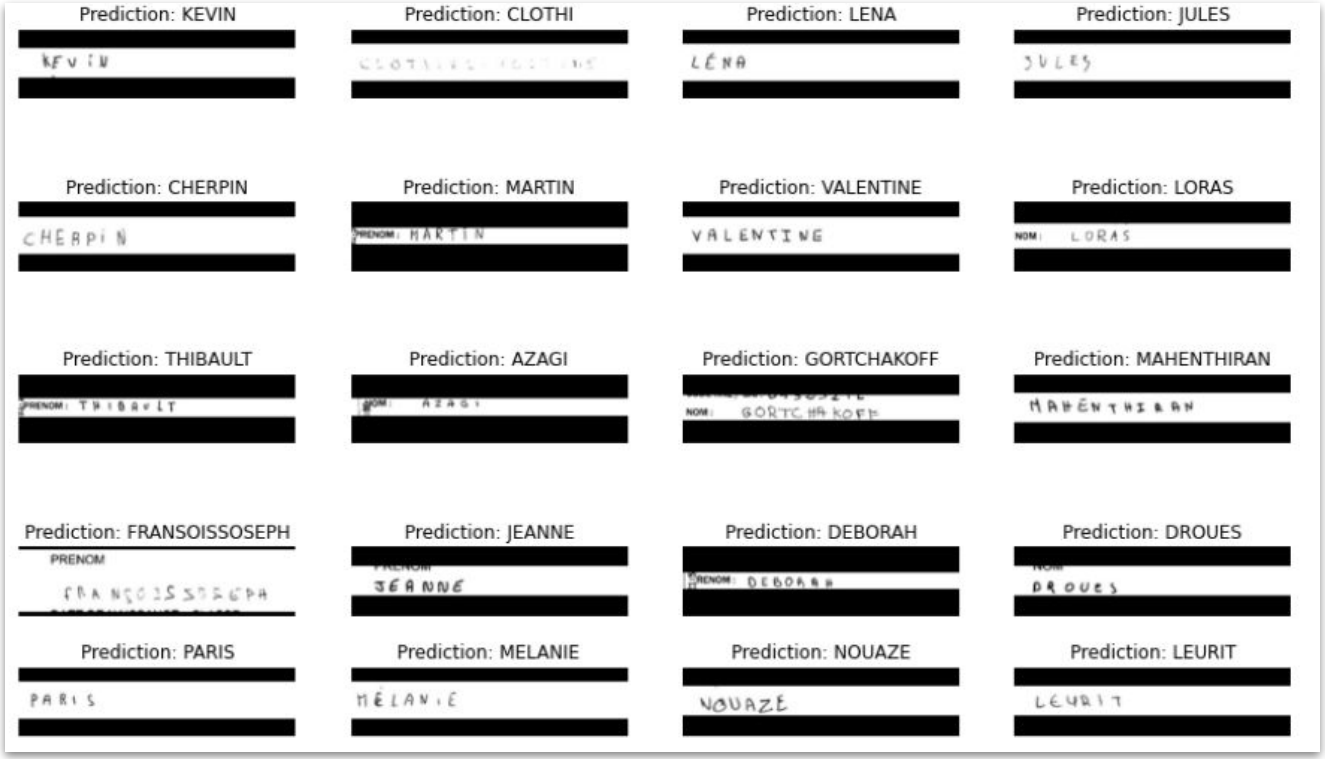
→ **Sampel 300 images -> sekitar 1-2% tidak dapat dibaca**

Orthogonalization Technique to Tune the Model

techready PRO Professional Academy digitalent KOMINFO

Model I	Model II	Model III
2 CNN layers & Max pooling	4 CNN layers & Max Pooling	4 CNN layers & Max Pooling — Dropout (0.2)
1 dense layers (64)	2 dense layers (256, 64)	2 dense layers (256, 64) — Dropout (0.2)
2 RNN layers (128, 64)	2 RNN layers (128, 64)	2 RNN layers (128, 64) — Dropout (0.25)
1 softmax layers	1 softmax layers	1 softmax layers
CTC	CTC	CTC
Bias Optimal: 1% CER on Training: 6.1% CER on Validation: 6.8%	Bias Optimal: 1% CER on Training: 2.7% CER on Validation: 5.7%	Bias Optimal: 1% CER on Training: 3% CER on Validation: 4%
Underfitting: Coba fokus untuk menurunkan bias di model II	**Overfitting:** Coba fokus untuk menurunkan variance di model II	**Slap untuk di test performanya pada test dataset**

Performa pada Test Data Set



	FILENAME	IDENTITY	PREDICTION	CER	WER
0	TEST_0001.jpg	KEVIN	KEVIN	0.0	0.0
1	TEST_0002.jpg	CLOTAIRE	CLOTHI	37.5	100.0
2	TEST_0003.jpg	LENA	LENA	0.0	0.0
3	TEST_0004.jpg	JULES	JULES	0.0	0.0
4	TEST_0005.jpg	CHERPIN	CHERPIN	0.0	0.0
...	...	...	...	...	...
3195	TEST_3203.jpg	ENZO	ENZO	0.0	0.0
3196	TEST_3204.jpg	VALENTIN	VALENTIN	0.0	0.0
3197	TEST_3205.jpg	DUMONT	DUMONT	0.0	0.0
3198	TEST_3206.jpg	JUSTINE	JUSTINE	0.0	0.0
3199	TEST_3207.jpg	BROUQUIER	BROUQUIER	0.0	0.0

3200 rows × 5 columns

CER_test_average: 4.27765584375

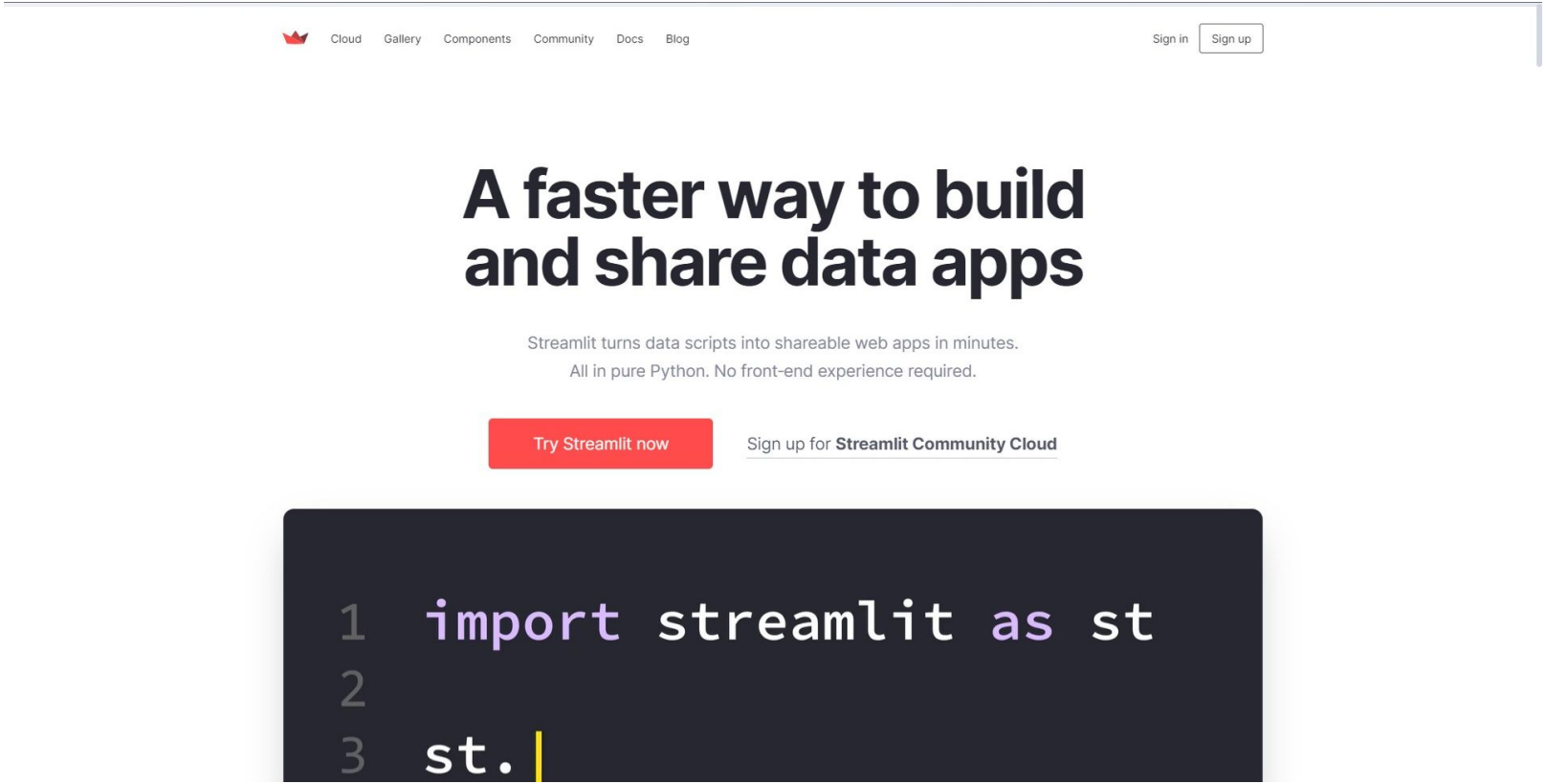
Komparasi: Good-> Printed Character (1-2%), Handwritten (<5%)

Error Analysis

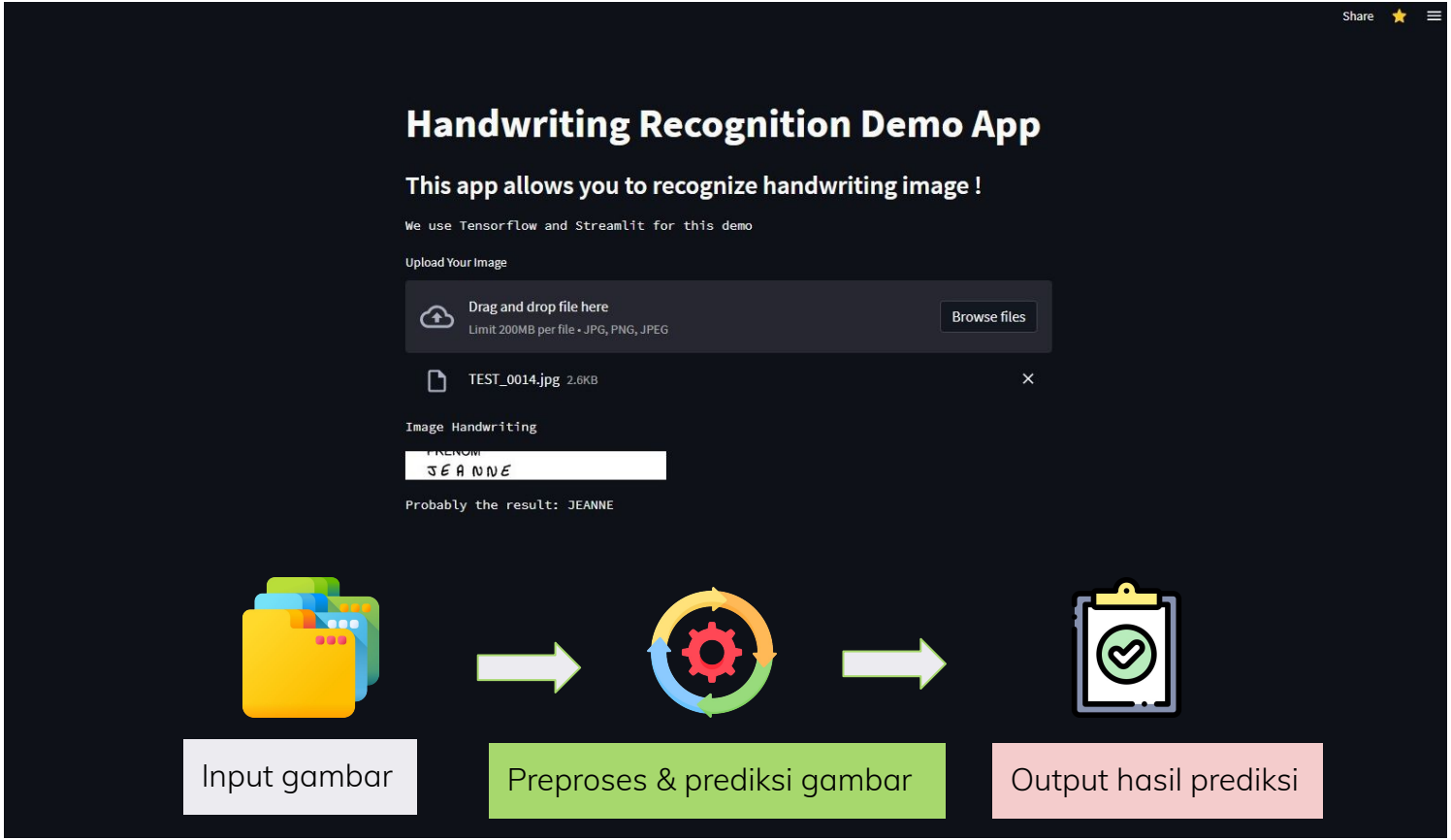
7. Deployment

techready PRO Professional Academy digitalent KOMINFO

Web Application



Web Application



Project Reference

Dataset : Data Handwriting
(https://drive.google.com/drive/folders/1izbmpOCkNmVudpUVOmiVV7zW6Pt511p1?usp=sharing**)**
Jurnal :

- **XinSheng, Z., Yu, W. 2022. Industrial character recognition based on improved CRNN in complex environments. Computers in Industry 142.** https://doi.org/10.1016/j.compind.2022.103732
- **Hao Zeng, 2020. An Off-line Handwriting Recognition Employing Tensorflow.** https://doi.org/10.1109/ICBAIE49996.2020.00040
- **Weibo Yu, Chunyu Guo, Keping Liu, Hongtao Yang. 2020. Handwritten Digital Recognition Optimization Method based on Deep Learning.** https://doi.org/10.1109/CAC51589.2020.9327647
- **SaraAqab, Muhammad Usman Tariq. 2020. Handwriting Recognition using Artificial Intelligence Neural Network and Image Processing.** https://doi.org/10.14569/IJACSA.2020.0110719
- **G.R. Hemanth, M. Jayasree, S. Keerthi Venii, P. Akshaya, and R. Saranya. 2021. CNN-RNN Based Handwritten Text Recognition.** https://doi.org/10.21917/ijsc.2021.0351

techready PRO Professional Academy digitalent KOMINFO

Thank You