



Customer Churn Prediction on Telcozee

**Neural Network
Approach**

McJager Consulting (MJC)

Project Summary

Customer Churn is one of the most common problems in business company where customers will continue to use their services or not. This problem need to avoid by Telcozee to prevent loss of customer, because *it is always more difficult and expensive to acquire a new customer than it is to retain a current paying customer.*

To respond these problem we (MJC) want to offer **an analysis of sleeping customer prediction system** which can predict the probability of churn customer based on 90 days of grace period. *Churn probability will be predicted based on data 45 days prior to contract expiration, in order for Telcozee to have sufficient lead time to affect customer behavior with an incentive offer.* We will solve this problem using **One of Neural Network Model (deep learning)** what that algorithm can produce high accuracy in a large-scale of data, and fast in build a model,

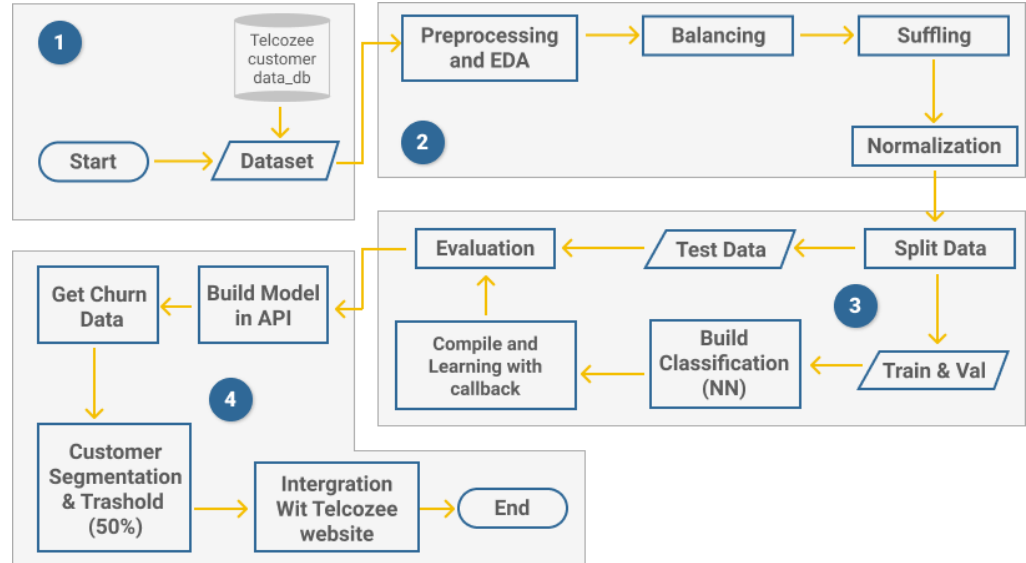


Technical Approach

Algorithm

- Take the sample of data from Telcozee database as our churn dataset.
- For each data, perform exploratory data analysis (EDA) and data cleansing such as missing value handling, data encoding, imbalance class handling, data shuffling and normalization.
- Split the data into train (with validation) and test (75% train; 25% test), then try to build a model using RNN, ANN and CNN. Evaluate the model and choose the best one.
- Build the classifier model as an API and get the churn customer. Make a customer segmentation model and take 50% data with high loyalty level.
- Try to integrate that model with Telcozee Website.

Divided Into 4 Main Stages



Technical Approach

1. Data Collection

Build a pipeline to get internal data source (from Telcozee) and try to get external data for analysis.



2. Data Preprocessing

Read the data using Pandas and try to explore the data using matplotlib and seaborn to get a lot of insights from that data. Make sure that we are done handling missing value, outlier, imbalance class, etc.



3. Build Model and Evaluation

Build model using Deep Learning algorithm like (RNN, CNN and ANN). We estimate that the accuracy of the model **can reach 80%**.



4. Deployment

We **will integrated the model result with Telcozee website**. Or we can build the different website specially for analyzing the customers behavior.



**4 Main Stage
For Development
Process**

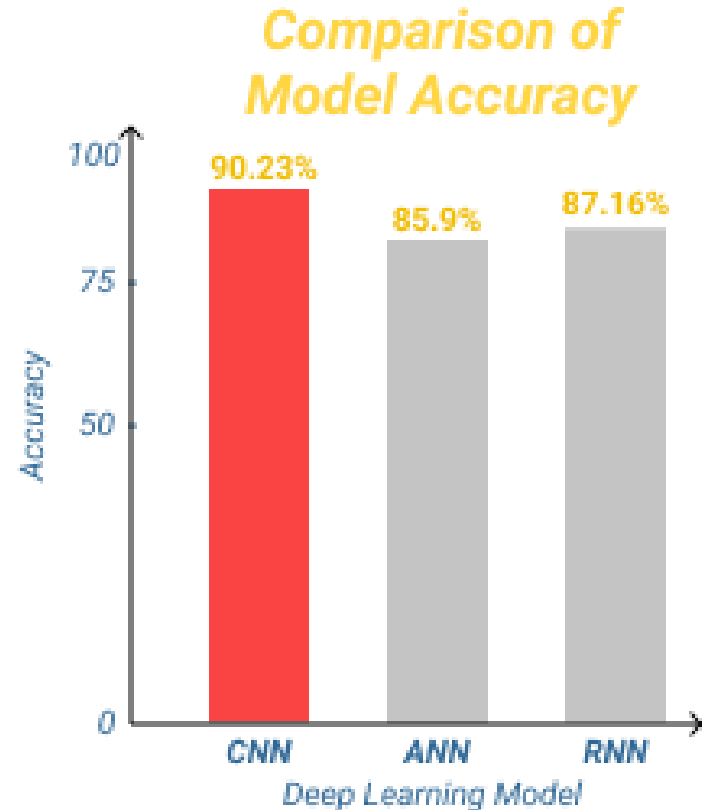
Data Sets

- We plan to work with the Telco dataset publicly provided by [kaggle](#) to see the Telco data variables in general and try to build our model before implement in Telcozee's data.
- We Plan to work with Telcozee's customer data that we get from the pipeline, try to get data with a balanced number of class. If it can not, at least the data difference is not too much.
- Cooperation with AFPI to get the customer behavior, especially in loan or investment. From that data we can see how reach are they, so we know their segment.

Name	Description	Length	Type	Classified
Kaggle Telco Data	See the telco data in general as reference	7043 rows x 21 columns	Customer Data	Yes: Yes/No
Telcozee's customer data	Real data for this project	We need about 500.000 data	Customer Data	Yes: Hand classified
AFPI Data	Get the loan and investment user data (see their loyalty)	Medium-Long	Customer Data	No

Experiment and Evaluation

- We also try to build the model in several deep learning algorithm such as RNN, ANN and CNN. Using **Telco kaggle dataset**, the best model accuracy we obtained when we use **CNN Model**. This is one of the reason why we suggest this model to solve customer churn in Telcozee.
- We believe that we can increase the model accuracy, remember that we **will collaborate with AFPI** to get the customer behavior data. And this data has a big impact for predict customer churn.



Software

Programming Languages

Build Model

Python 3.8 and Library Pandas, Matplotlib, Seaborn, scikit learn and Django.

PostgreSQL for database

Build Website

React Js, Node Js



Evaluation Software

Model Evaluation

For training model we used K-fold cross validation from sklearn library.

For testing model we use confusion matrix from sklearn matrix confusion, then plot it like heat map.

Website Evaluation

We will test the website performance using gtmatrix and wabepagetest (online checking).



Collaboration Software

Check Our Progress

As a partner you can check our progress through **Jira** and we can communicate through **slack**.

Try our model

We will send you our API url after the model already done. You can check the model using **Postman**. And you can check our beta website after it's already done. We will send the url soon and the guide book.



Milestones



Week 1-2

Business Understanding, Data understanding, data collecting, and pipelining



Week 3-5

Exploratory data analysis and data preprocessing



Week 6-7

Build model, make API, and evaluation together with mitra



Week 8

Integrete the model (API) with Telcozee's website



Week 9-10

Write a report, figure and final presentation materials

*We need 2.5 months to finish this project, but this time is so flexible (can finish faster).



THANKS!

DOES ANYONE HAVE ANY QUESTIONS?

Mjcconsultan_indo@mjc.com

+62 85 2582 86791

www.mjc_consultan.com



MJC_Indonesia



mjc_Indonesia



mjc consultan indo



mjc_Indonesia