

MY PORTFOLIO

(Last update: April 2020)

<https://dindatasi.github.io> | <https://github.com/dindatasi>

SELECTED PROJECT HIGHLIGHTS

Ongoing personal projects

I'm casually working on several personal data analytics projects on my free time:

1. Python Chart Gallery (https://github.com/dindatasi/python_charts) - a collection of charts with complete attributes (legend, label, etc). The goal is to reduce time needed to draw a nice and easy to read charts in Python since the default viz from library are not normally user-friendly
2. My Spotify (https://github.com/dindatasi/my_spotify) - fun project using Spotify API to play around with my Spotify user data. The plan is to make several interesting viz and stories later out of this.
3. My personal page (<https://www.dindatc.com/>) channeling my old hobby of web design.

Project highlights at trivago (Marketing Analytics, 2019-now)

- Improvement of cross-device measurement: used third party analytics service to match users and analyze their booking journey & cross-device spill across channels within the attribution period, and then calibrate the cross-device calculation model. In addition, we also ran several experiments to validate the existing model & assumptions.
- Paid traffics value predictions: researched, developed, and evaluated multiple models (4PL Regression, decision tree regression, and existing booking prediction algorithms) to predict value of sessions and maximize return on investment in paid marketing channels.
- Various exploratory analysis to understand customer acquisition, behavior, and retention.

Project highlights at GO-JEK (2018-2019)

- Measurement & evaluation of marketing campaigns on partner e-commerce sites
- Merchant targeting for POS products.
- Exploratory analysis & customer profiling (digital wallet product): identify journey from acquisition, cross-sell & up-sell opportunities, and segmentation.
- Extended quantitative analysis result & figured out the unknown by collaborating with qualitative research teams
- Market size & our market share estimation in e-commerce payment industry. The result was very close the subsequent extensive research done by external research firm.

Next To Visit- Destination Recommendation System (2018)

https://github.com/dindatasi/next_to_visit

I collected thousands of place of interests and restaurants data in London from Trip Advisor to build a web-based recommendation system. User can input a destination, and the system will recommend top 5 similar places based on similarity, popularity, and distance from the location. The web app was built with Python and Flask, and the libraries used to build the model are Scikit-Learn and NLTK.

Dissertation Project With Warner Music Group (Marketing Analytics)

This is a project for my dissertation with Warner Music Group UK. The research used quasi-experiment concept with difference-in-difference regression to estimate the effect of playlist as a stream source. For this project I am using Stata & Python, with data science & visualization libraries such as pandas, numpy, bokeh, plotly.

COURSEWORK PROJECTS (MSc Business Analytics at UCL)

Warner Music: Predicting The Success of Artists On Spotify

This is a group coursework for Predictive Analytics module. The case & dataset were taken from last year's dissertation project with Warner Music. In this project we used around 3.8 million rows (10% of the original dataset) of Spotify streaming data to determine whether an artist will be successful based on the criteria given by Warner Music's analyst. The main challenges in this project include imbalanced dataset (only 13.6% of the artists on the dataset are successful) and handling lots of categorical variables. We tried Logistic Regression, KNN, Random Forest, and several ensemble models to find the best predictive models. Our final model with Random Forest Classifier successfully achieved ROC AUC of 89.4%. Some of the tools we used include Python, Pandas, Scikit-learn, SQL, etc.

Airbnb Superhost Evaluation

[Check the slides here](#) and [the Tableau Dashboard here](#)

As part of Business Strategy & Analytics Individual Coursework, I analyzed the consistency of ratings received by hosts who have more than one listing, and then evaluated whether the *Superhost* criteria should be changed. For this analysis I used Python, Jupyter Notebook, SQL, and Tableau.

Kaggle Challenge: StumbleUpon Evergreen Classification

[Check code here](#) and [Kaggle page here](#)

Participated in past Kaggle challenge as part of individual coursework. I built a model to classify evergreen and ephemeral contents for Stumble Upon. My final model is Logistic Regression with 88% ROC AUC. The model was built in Python using scikit-learn. In this project I applied various NLP techniques such as TF-IDF, stemming & lemmatization, and sentiment analysis.

Mosey

https://dindatisi.github.io/mosey_landing/

This is part of Business Strategy & Analytics Group Coursework. Mosey is a navigation app that provides choice for safest route instead of only the fastest route. We took crime data from data.police.uk and develop an algorithm to recommend the safest route for walking. The gamification and place of interests suggestions offered also encourage people to walk more instead of using other modes of transportation. Aside from the product roadmap formulation, my main responsibility was to develop the initial routing algorithm, integrate with GoogleMaps and Instagram API to get the list of place of interests, as well as to develop the landing page for the app. The tools used in this project include Python (mainly Pandas & NetworkX) and Tableau.

Amazon Fake Review Detection

https://github.com/dindatisi/MSING055_Amazon_Review

<https://github.com/Maytudboon/MSING055>

This is my group's final project for Programming for Business Analytics module. The idea was to use network analysis and clustering technique for anomaly detection. We scraped product review data from amazon by following snowball method. Then we constructed and compared the network, with *profile id* as the node, which will be connected to each other if two *profile ids* have reviewed more than two products in common. We found huge discrepancy in which the network for the suspected fake review group was very crowded. Then we extracted the list of suspects and matched it with the reviews clusters. The clusters were created using simple K-means with 5 clusters. We found two clusters contained >77% of reviews from our the fake reviewer suspect. The tools I used in this project are Python, Jupyter Notebook, Weka, and MySQL.

OTHER DATA PROJECTS

Analysis published on medium

- Ruangguru vs Zenius: <https://medium.com/@dindatc/is-ruangguru-more-popular-than-zenius-2c5531c36d12>

- Jakarta Flood 2020 & Anies Baswedan: A Perspective Based on Google Search Trend: <https://medium.com/@dindatc/jakarta-flood-2020-anies-baswedan-a-perspective-based-on-google-search-trend-ac505a0584b2>

WILD Product Metrics Dashboard (2018)

I developed a dashboard to track product metrics for WILD Technologies AI. The dashboard help them to identify not only the growth of their product, but also their top user segment and enable them to pick the most relevant customers for user interview. Tools used during the development include Metabase, PostgreSQL, and Python.

3D Ring Visualization (2017)

An interactive visualization of Midtrans transactions ring in 3D World, this desktop app enables user to explore the 3D world by using Leapmotion. The app read the transaction data in JSON file, and then generates the network graph/ring that consist of nodes. A node can represent buyer's email, phone number, or card number. By looking at the visualization, user can point out fraudsters easily.

Mycroft - Merchant Evaluation Portal (2017)

A portal to help with merchant onboarding decisions by automatically crawling and evaluating their website and social media accounts using LSA & sentiment analysis, and then provide recommendation whether to approve or reject the merchant.

Machine Learning for Aegis Fraud Detection System (2016-2017) <http://midtrans.com/aegis>

Aegis is one of the most sophisticated fraud detection system in Indonesia. It combines pre-determined rules created by human with a machine learning model that can catch fraudulent transactions. The model provides transaction fraud score and signal (reasons behind the score) which was implemented into Fraud Detection System, with ~75% accuracy and <40ms average response time despite handling large transaction volume.

Targeted Promo Analytics (2017)

This is a hackathon project, in which I worked in a team of five. This web app enables e-commerce and banks to effectively target specific customer segments for promotion, and design a promo that will best suit them based on the predictive model. User can submit the promo specifications, and then get estimated market size and the likelihood of using the promo for each of their customers. The app became the first hackathon product to be marketed by Midtrans.

Solutions Dashboard (2017)

The project started as an initiative to encourage data-driven decision-making within the company. We developed a dashboard using Tableau to provide insights based on sales and marketing stats. By using this dashboard, sales team's leaders can also easily track individual and team's performance according to their KPIs.

Anomaly Detection for The Biggest Private Bank in Indonesia (2016)

A proof of concept to analyze internet banking transactions data and discover the anomalies. Using big data tools & technique, we successfully identified suspicious and anomalous customers. We presented the result to senior managements of the IT Division of the Bank, and they decided to extend the proof of concept into full project.

WEB & MOBILE APP PROJECTS

Promo Engine (2017) <https://api-docs.midtrans.com/#credit-card-promo-beta->

Scrum Master

Promo Engine API serve as additional services to Midtrans Core API. It allows merchants to use certain promotions for credit card transactions to be offered to their users. We also provide dashboard to easily add and manage promotions.

Payment Landing Page Development for One of The Biggest Airlines Indonesia (2016)

Project Manager

Developed payment landing page that process IDR billions of transactions daily. We were integrating more ticket retail channels owned by the airlines to our payment landing page. We worked with their other vendors as well, including their other web developers and the booking engine providers.

Docspace.id (2016) <http://docspace.id/>

Product Consultant

This is a collaboration between my tech consulting firm and two doctors. The idea is to create a "space" for doctors to get various job-related information, including vacancies and events. We were highly involved in the market research, product planning, and development process. The pitchdeck we created also help Docspace to win an award at a bootcamp.

Midtrans Mobile SDK (2016-2017) <https://mobile-docs.midtrans.com>

Scrum Master/PM

SDK to let merchants implement Midtrans Payment Gateway seamlessly in their native iOS and Android app. This SDK carries all features offered by Midtrans Payment API, as well as giving merchants the choice to implement the drop-in UI kit (for quicker integration) or the core kit (API-only implementation, for more customization). During this time, I was also in charge to help several key merchants to get the most out of our Mobile SDK.

BBM Pay (2015-2016)

Scrum Master

A responsive web payment page for Blackberry Messenger. We worked directly with the Senior Manager of Blackberry Messenger to drive more sellers to use BBM Channel to sell their products, and then let their customers pay using the BBM Pay payment page. Aside from the payment page, we also provided the payment link generator and a portal to manage inventories & record transactions.

SIEndo - Endowment Fund Information System & SIMon - Commercial Unit Monitoring Information System (2015)

System Analyst

Hired by BPUDL ITB to develop both SIEndo & SIMon. SIEndo give easy access for donors to find information related to ITB's endowment fund as well as their donations to ITB. Meanwhile, SIMon was planned to track progress and performance, such as financial KPIs achievements, of 14 commercial units under BPUDL ITB. The web portals was developed using laravel & PHP, and we worked directly with the Head & Secretary General of the organization to gather requirements.

e-SDLC (2014)

Software Developer

Developed a web-based IT project management application to submit project proposal & track the development progress. The app enable users to add/request new IT project, assign employees to projects according to their departments, and notify users about new project assignments. The purpose was to promote better visibility & transparency of project's progress, as well as to get rid of paper-based documentations. I developed it with two other interns at Bank Rakyat Indonesia, one of the biggest state-owned bank, and it became one of a few apps developed by intern that were being deployed to the central office.