

Hyunjin NAM

Data Analyst

[in linkedin.com/in/hyunjin-nam-5bb44a152](https://www.linkedin.com/in/hyunjin-nam-5bb44a152) hyunjinnam.com
☎ +46 70 027 64 18 @ jinanam0116@gmail.com
📍 Stockholm, Sweden 🇰🇷 South Korean



I am a Data Analyst with a background in statistics, working mainly with Python and R. I specialize in machine learning applications with focus on customer base analysis. By using a statistical approach, I am building a predictive models to help businesses understand customer needs better.

EDUCATIONS

- 2019 Master of Statistics | Uppsala University
- 2017 Bachelor of Statistics | University of Seoul
- 2017 Bachelor of Social Welfare | University of Seoul

PROFESSIONAL EXPERIENCE

- | | |
|---|--|
| November 2019
Now | Data Analyst Tele2, STOCKHOLM, Sweden <ul style="list-style-type: none">➢ Tele2 AB is a European telecommunications operator headquartered in Stockholm, Sweden.➢ Building a machine learning model targeting the right customer to cross-sell, up-sell and retention (cross-sell broadband to mobile customers, up-sell TV add-ons to TV customers).➢ Building a Tableau dashboard to show a real-time overview of customers' behaviours to gain business insight (the volume of the customers' propensity score for each event, 5G data consumption).➢ Analytical support to the marketing team by following up the effectiveness of personalized offers using A/B testing, significance testing➢ Ad-hoc analysis and regular reporting using SQL <div>Python R Tableau SQL Snowflake DBT</div> |
| July 2020
December 2020 | Course instructor Hansung University, SEOUL, South Korea <ul style="list-style-type: none">➢ Taught a statistics course "Basic Statistics and Data Science with R" at Hansung University. The course is consists of 7 online lectures including laboratory sessions with R.➢ Trained students to be able to step into the data science world dealing with real data➢ Introduced essential statistics skills such as data handling (Dplyr), visualization (Ggplot2), T-test, and logistic regression (GLM). <div>Statistics R Dplyr Ggplot2 GLM</div> |
| January 2019
November 2019 | AI Developer BYON8, STOCKHOLM, Sweden <ul style="list-style-type: none">➢ Byon8 is a health tech company using artificial intelligence in medical diagnostics. Byon8 provides diagnostic suggestions from the AI engine to ensure secure diagnoses, recommendations and treatments.➢ Responsible for data preprocessing, data mining with demographic and clinical data.➢ Built a tree-based model that can help diagnose diabetes. <div>Python R Numpy Scikit-learn Pandas Seaborn</div> |
| September 2017
Jun 2019 | Digital Ambassador Svenska Institutet, STOCKHOLM, Sweden <ul style="list-style-type: none">➢ Studyinsweden is an organization built and maintained by the Swedish Institute (svenska institutet) and is tasked with promoting Sweden for prospective international students abroad.➢ Provided contents regarding the Swedish higher education system and living as an international student in Sweden➢ Responsible for writing and managing the official instagram @Studyinsweden. <div>Marketing</div> |

</> SKILLS

Programming	Python, R, SAS
Databases	SQL, Snowflake
Data analysis	NumPy, Pandas, Scikit-learn, Dplyr, Ggplot2, GLM
Others	Vuepress, HTML5, CSS, jQuery

📁 PROJECTS

RECOMMENDATION ENGINE

2021 - 2022

Tele2, Sweden

I have been working on building a recommendation engine that suggests the following best action to an existing customer providing the personalized offer. I have worked explicitly on building cross-sell TV, BB and TV add-on models. Especially when Tele2(A mobile-based telecommunication company) merged with Com Hem(A TV/BB based telecommunication company), we required a better solution to cross-sell TV, BB to the existing mobile customer and vice versa.

Predictive models can help distinguish the right target for personalized offers more reliably by combining the various customers' characteristics. The functional features are regarding customer's subscription(Binding, list price, net price), household data(household size, the number of children, Living conditions, infrastructure), and the customers' behaviour (Data consumption, the number of calls, TV viewing duration, TV viewing channels).

The models are based on Random Forest, giving the output probability of engaging in an event at a particular time. I used model metrics (AUC, ROC, precision, recall, feature importance) to select features and tune parameters.

The output from the predictive models is combined with data on the impact a successful campaign will have on the customer lifetime and recurring value. By prioritizing the customer lifetime impact, We send a personalized offer to the existing customer through a 1 :1 marketing channel.

[Git](#) [SQL](#) [Snowflake](#) [Docker](#) [Python](#) [Scikit-learn](#) [Pandas](#) [Matplotlib](#) [Numpy](#)

CUSTOMER BASE OVERVIEW

2019 - 2022

Tele2, Sweden

The customer base overview is the project I am driving as a data analyst in the Commercial Analytics team, Tele2. I am making a report that can give our colleagues a summary view of our customer base and a chance to get a deeper insight into what happened during the past year. The customer base overview project is an excellent opportunity for employees to learn an overall view of the customer base and how the development had proceeded.

I have been working on the project for three years, and traditionally the data extraction is done on SQL and visualization and explanation had been shared through the PPT presentation. For the reproducible purpose, I initiate using the Tableau dashboard to share updated data in real-time, and everyone can have access to the overview at the end of the year and whenever they need it. Using the Tableau dashboard made our team's workload less as our stakeholders can actively go into the dashboard and search the updated status by themselves, and it is more manageable within the group to share the query that can reproduce equivalent output.

The contents include valuable information about our customers' demographics(Region, age), Stock of the base(The total number of subscriptions/customer/HH, product type), Value(List price/ Net price, Family subscription, Fixed and mobile conversion), Usage(5G data consumption, TV viewing duration, linear channels), Sales(Sales channel, mobile hardware, cross-sell/ New intake) and Churn(The number of churned subscription, portation, revenue loss)

[SQL](#) [Snowflake](#) [Dbt](#) [Tableau](#) [Powerpoint](#)

Uppsala University, Department of Statistics, Independent thesis Advanced level (degree of Master (Two Years))

[Download](#) github.com/hyunjin-nam/thesis-diabetes

The aim of this study is to develop a statistical model to predict type 2 diabetes based on the tree-based model. Furthermore, the aim to compare classification with current medical criteria. Decision Tree, Random Forest, Boosting with a XGBoost algorithm is used as a classification method to predict diabetes. The results show XGBoost outperformed the two other models in yielding highest classification rate, with a 84.6 percent test accuracy. Two interesting findings from this paper are : 1) Ensemble methods such as Random forest and boosting can be easily overfitted on training data, but this problem can be solved with correct hyper-parameter tuning. And 2) Tree-based methods such as XGboost and Random Forest can solve variables' multicollinearity problems.

Machine learning Medical health R Python

LANGUAGES

Korean	●	●	●	●	●
English	●	●	●	●	○
Swedish	●	○	○	○	○