UC San Diego

Master of Data Science program

"To fulfill my goal of becoming a global solution provider"

Written by Minseok Oh
Dec, 2022

Prepared for
University of California San Diego

University of California San Diego

CONTENTS

My approach & Understanding

1. Self Introduction (Personal)
2. Self Introduction (Professional)
3. My understanding of Master of Data Science program
4. Why UCSD ? (My progress)
5. My course schedule in UCSD
6. My research and area of interest
7. Why UCSD ? (My direction of research)

To be CLEAR… Self Introduction (Personal)

- ✓ **Understand how to collaborate and share things with each other than staying in individual happiness Experience the level of creativity allowed in Canada, I had a dream of going abroad, on a bigger stage.**



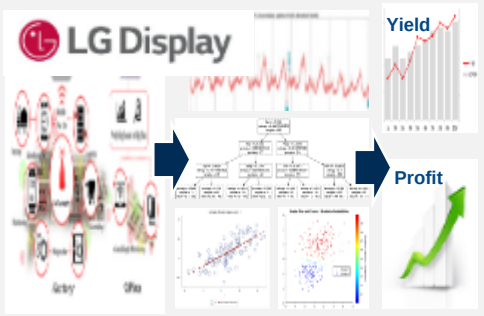
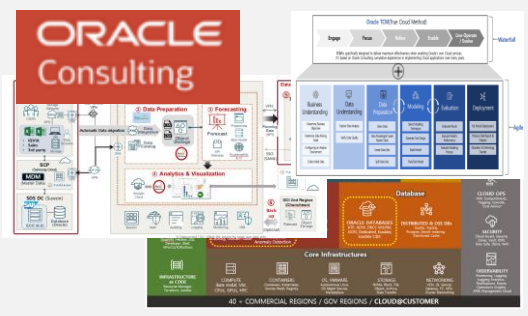


- ✓ Have been grateful and satisfied with all the things given to me - a stable family, faith and love, and finan cial leisure that does not overflow. If I hadn't practiced sharing, such as IT Volunteer work in Ecuador and supporting students in Zimbabwe, I might have lived a very personal life. However, through such a preciou s experience, I hoped for a shared future rather than staying in individual happiness.
- ✓ When I was in Canada, I saw firsthand the degree of diversity accepted and appreciated in Canada and the level of creativity allowed. I was moved by the possibilities that an open society can bring to its members. Since then, I had a dream of going abroad, on a bigger stage. To this end, I made endless efforts, such as a Master degree program and studying English, and came to Oracle, which would become a stepping stone to my future dream. At Oracle, I had a lot of experience, such as carrying out overseas projects with global experts.
- ✓ Now is the time to take the next step toward my dream. My next destination is UCSD, and I want to work at a global tech company for 3 years, understand the flow of the US and global markets, and start my own business in 5 years.

To be CLEAR… Self Introduction (Professional)

✓ **Ready to share my experiences in the field of data science, and be confident to create synergy in research activities and collaborate with colleagues at the University of California San Diego**

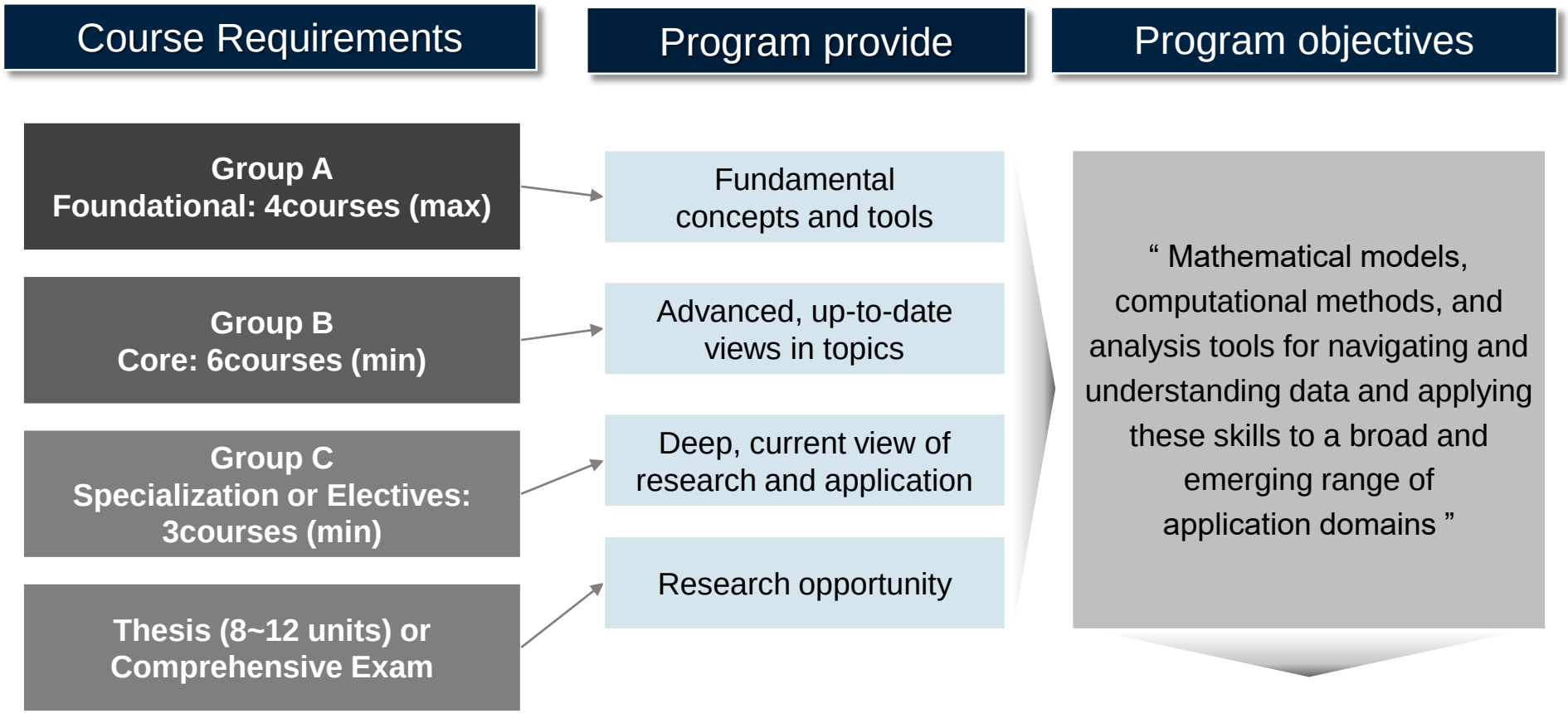
My expertise in Data Science field and my enthusiasm for learning

• Data Scientist (LG Display) ('13.01~'19.05)	• KAIST Master of Science in Information Management ('16.09~'19.01)	• AI/ML & Data platform Consultant (Oracle) ('19.05~present)	• Enthusiasm for learning (~Continuous)
" Provide business value by analyzing data and developing models so as to excavate insights for the front-line employees "	**" An interdisciplinary program that combines data analysis, strategy, and management (Lectures: Business Analytics, Big data Analysis, App development, IT system Design, IT business Management) "**	**" Provide consulting services to design data architecture and develop AI/ML & Data platform to build data-driven enterprise "**	**" Externally activities to improve my te chnical knowledge in Data Science and academic subjects"**
			
• **Implement multiple ML/data analysis projects** → Develop an anomaly detection and monitoring system with ANN algorism, enabling early error detection of pixel deposition facilities. (Reduced annual defect rate by 0.2%) ▶ Python, SAS(JMP), Spotfire, Matlab, Excel	• **Research:** Expected Values on the Continuous Intention to Use IoT Products from the Perspective of Expectation-Confirmation Theory (user, survey data) • **Data analysis & ML Projects** with startup companies (Recommendation system, churn &conversion rate, target marketing, and log data analysis) ▶ Python, SPSS, Smart PLS, MySQL OracleDB	• **Implemented multiple AI/ML & data platform projects** (Industry: pharmaceutical, distribution, manufacturing, insurance, banking) • **Engage in proof-of-concept work or product demonstration** for clients to deliver data strategy and insight based on clients' requirements. ▶ Python, Oracle solution (OML, OAC, Data Science, AI Service)	• **External efforts:** **Oracle professional training** (Analytics, Database, Cloud Infrastructure, Architecture) **Courses** (Linear Algebra, Multivariate Calculus, Statistical Inference, Randomized Algorithms, Data Structure) **Conference & Volunteer work** (KORMS conference, NVIDIA GTC 2018, IT Education volunteer(Ecuador))

To be CLEAR… My understanding of Master of Data Science program at UCSD

✓ **Master of Data Science program offers knowledge and skills to be successful at performing data-driven tasks, and lay the foundation for future researchers who can expand their knowledge in Data Science.**

UCSD's Master of Data Science Program overview



Goal: **" Boost knowledge of how to apply these skills to real-world problems "**

To be CLEAR… Why UCSD Master of Data Science program ? (My Progress)

- ✓ **Having master's degree in UCSD which is the best Data Science program across the globe is prerequisite for my future objective to start my own business**

- ✓ In 2019, I decided to pursue my journey at Oracle in order to aid in preliminary work for companies in a variety of industries seeking to start data transformation. Oracle allows me to develop into a better-equipped analyst and entrepreneur. I'm able to understand different market demands by facing unmet needs by serving consulting services to multiple clients in a diversity of industries. I plan to develop my own solution for the existing numerous Small and Midsize Businesses (SMB) each with its unique set of needs. With the lessons from my career and degree program, I aspire to develop AI and ML-based solution packages that will help SMBs in their operations across the world.

- ✓ Having a master's degree from UCSD which is the best Data Science program across the globe is a prerequisite for my future objective to start my own business. I will make my partners and, expand my technical and theoretical knowledge. Also, I will have time to quench my thirst by reading and understanding various research papers and finding things that can be applied. Moreover, I will publish a research paper within my interest to deepen my knowledge. I'm certain that I will be able to accomplish my goals and create meaningful results by partaking in the Master of Data Science program at UCSD.

To be CLEAR… My course schedule in UCSD

✓ **Quench my thirst through Core and Specialization coursework (In particular★) to fulfill my future objective**
Plan to take 5 quarters (52 credits): Intro (8), Core (24), and Specialization (8) and select thesis option (12)

	Year 1			Year 2	
Calendar	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Quarter 5
Credit	12 credits	12 credits	12Credit	12 credits	4 Credit
1	DSC 211: Introduction of Optimization	DSC 241: ★ Statistical Models	DSC 206: ★ Algorithms for Data Science	DSC 298: Independent Research	DSC 298: Independent Research
2	DSC 210: Numerical Linear Algebra	DSC 204B: Big Data Analytics and Applications	MGTA 455: Customer ★ Analytics	DSC 244: ★ Large-Scale Statistical Analysis	
3	DSC 240: Machine Learning	DSC 242: High- ★ dimensional Probability and Statistics	DSC 298: Independent Research	MGTA 479: Pricing ★ Analytics	
4					

★ Important courses for me to enhance my knowledge are needed

Curriculum		
Option (Exam or Thesis)	Thesis Option: Berk Ustun, Julian McAuley	
Elective and Specialization Courses (Marketing)	Group C: Marketing: 2 courses	Electives: 5 courses
Core Courses	Group B: Core Courses 6 courses	
Introductory Courses	Group A: Introductory courses 2 courses	
Total	52 credits	

※ Graduate Requirement: 48units
40units (graduate-Lv), 2 of 10 are not directly related to DS

To be CLEAR… My research and area of interest

- ✓ **Would like to work on developing methods to promote the use of machine learning, especially with regard to the recommender system, customer behavior analysis, and forecasting.**

My interest in Data Science

- ✓ Recommendation model
- ✓ Time-series analysis
- ✓ Sales & demand forecasting
- ✓ Customer segmentation
- ✓ Anomaly detection
- ✓ Customer satisfaction prediction
- ✓ Social media analysis

Deepen my knowledge at UCSD

- ✓ Explainable recommendation
- ✓ Graph-based recommendation
- ✓ Sentiment-based recommendation
- ✓ Media (cartoon, music, book) recommendation
- ✓ Fake review detection
- ✓ High-dimensional Time-series analysis
- ✓ Forecasting for new launching products (Without historical data)
- ✓ Development of driver factors (Find out factors affecting the model's performance)
- ✓ Continuous decision-making under uncertainty

To be CLEAR… Why UCSD Master of Data Science program ? (My direction of research)

✓ **Deeply impressed by Berk Ustun and Julian McAuley's research, I would like to research with them capturing crucial factors that affect the performance of the recommender system and how to resolve them**

Based On Experiences

Description

LG Display
Data Scientist

" Sharpened my skillset in Data Science filed "
- Validate factors by using statistical methods
- **Development of prediction models and monitoring systems**
- Opportunities to analyze big data

KAIST
MS in Information Management

" Broaden my standpoints in Data Science filed "
- **Work with various industry startup companies**
- **Experience multiple types of data, recommendation system**
- **Publish research paper (Statistical methodology)**
- Understand Data-driven strategy

Oracle
AI/ML Data platform consultant

"Apply to real cases & Experience end-to-end service"
- Provide data management strategy for clients
- **Implement AI/ML, data platform consulting projects (Market Basket Analysis & etc.)**
- Understand architecture with Oracle's strong data product lineup

External efforts
course, training

" Enhance academic & technical knowledge "
- **Attend several mathematical and data coursework**
- Take professional training courses and earn certificates
- Attend academic & Global AI conference

Desire to work with faculty at UCSD

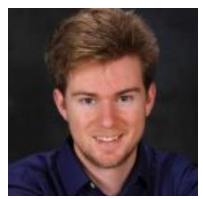
Berk Ustun

Rank List Sensitivity of Recommender Systems to Interaction Perturbations

→ Measures how **rank lists** generated by a given recommender system at test time change as a result of a perturbation in the training data

When Personalization Harms: Reconsidering the Use of Group Attributes in Prediction

→ Approach to personalization using group attributes such as gender, age, and blood type does not improve the performance of all groups. Discuss how this effect harms.



Julian McAuley

Query-Aware Sequential Recommendation

→ Among interaction data sets, important queries are ignored in Sequential recommenders

End of Document