

TING-YI (LISA) HUNG

tingyi.hung@utexas.edu • [linkedin.com/in/ting-yi-hung](https://www.linkedin.com/in/ting-yi-hung) • (737) 529-7321

EDUCATION

The University of Texas at Austin	Master of Science in Information Technology and Management GPA: 3.43/4	May 2019
Coursework Includes: User Generated Content Analytics, Big Data & Distributed Programming, Business Data Science, Data Management, Cognitive Computing, Advanced Data Mining & Web Analytics		
National Sun Yat-sen University	Master of Business Administration, Information Management GPA: 4.19/4.3	Jul. 2017

SKILLS

-
- **Tools/Cloud Platform:** Azure Data Factory, Azure Databricks, Azure DevOps, Trello, Jira, Oracle SQL Developer, Amazon EC2, Paperspace Fast.ai, Git, GitHub, Tableau, MS Word, Excel, PowerPoint, Google App Engine, Android Studio, Web Scraper
 - **Programming Languages:** SQL, Python (pandas, NumPy, SciPy, matplotlib, seaborn), Java, HTML, Kotlin

EXPERIENCE

Dell – Data Engineer Capstone Project Intern	Dec. 2018 – May 2019
<ul style="list-style-type: none">• Collaborated with a team of Data Architects, Software Engineers, and Data Scientists on the development of a new Data Ecosystem across Dell’s business units• Developed a standard/method for landing a specific data stream in the form of raw XML files by processing and structuring the data using Apache Spark framework within Azure Data Bricks• Evaluated and built proofs of concept for a data pipeline from On-Premise environments to Azure Cloud environment using Azure Data Factory to demonstrate that the same data stream can be leveraged by different teams and business units	
Digiwin Software – Technical Service Engineer	Feb. 2014 – Apr. 2015
<ul style="list-style-type: none">• Resolved customers’ problems, 340 cases per month, related to SmartERP system, including correcting abnormal data, dealing with compatibility between ERP and surrounding devices• Tuned operations of ERP to fit customers’ business processes; customers’ average satisfaction score: 9/10• Utilized SQL to correct abnormal data in SQL Server to improve the accuracy of customer’s financial statements• Recovered customers’ data when system anomaly occurred and created database backup jobs in SQL Server to maintain data durability	

PROJECTS

Movie Recommender git.io/fh6pA	2018
<ul style="list-style-type: none">• Built a movie recommender employing a data preprocessing pipeline in Apache Spark to transform dataset contained 100K+ ratings and 3,683 tag applications from MovieLens website• Graded by professor with the highest score in the class; proposed personalized movies based on genre, tags and ratings	
Analyzing BestBuy Reviews Project git.io/fh6hd	2018
<ul style="list-style-type: none">• Developed web scraping program automatically utilizing Python Selenium and extracted 3.5K reviews on BestBuy website• Executed data cleaning utilizing Natural Language Toolkit (NLTK) and performed sentiment analysis utilizing Python TextBlob Library to analyze the reviews• Analyzed data employing Latent Dirichlet allocation (LDA) topic modeling and visualized the results using pyLDAvis package	
Kaggle Competition: Find the Secret Binary Outcome	2018
<ul style="list-style-type: none">• Executed feature engineering and compared different models on the dataset where all the columns’ names were not given• Applied GridSearchCV to find the best combination of the parameters in XGBRegressor and combined with other model and tuned the weights between them; improved the accuracy from 0.84 to 0.89	
New York Stock Price Prediction	2019
<ul style="list-style-type: none">• Predicted stock price trends employing Auto ARIMA and LSTM in Python on historical stock prices of 10 companies from 2010 to 2015 and evaluated each model using RMSE and data visualization• Performed data cleaning and feature selection and applied ordinary least square model on firms’ financial indexes dataset• Assisted the stockholders in making decisions on when to trade stocks based on influential features	