

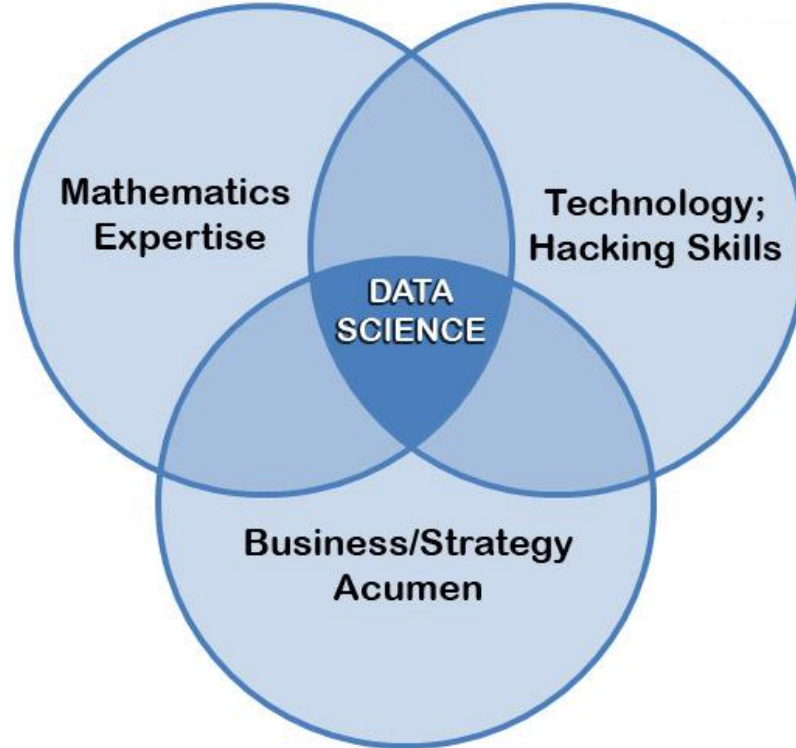
COMPETITION DETAILS

Dr Pan Yaozhang

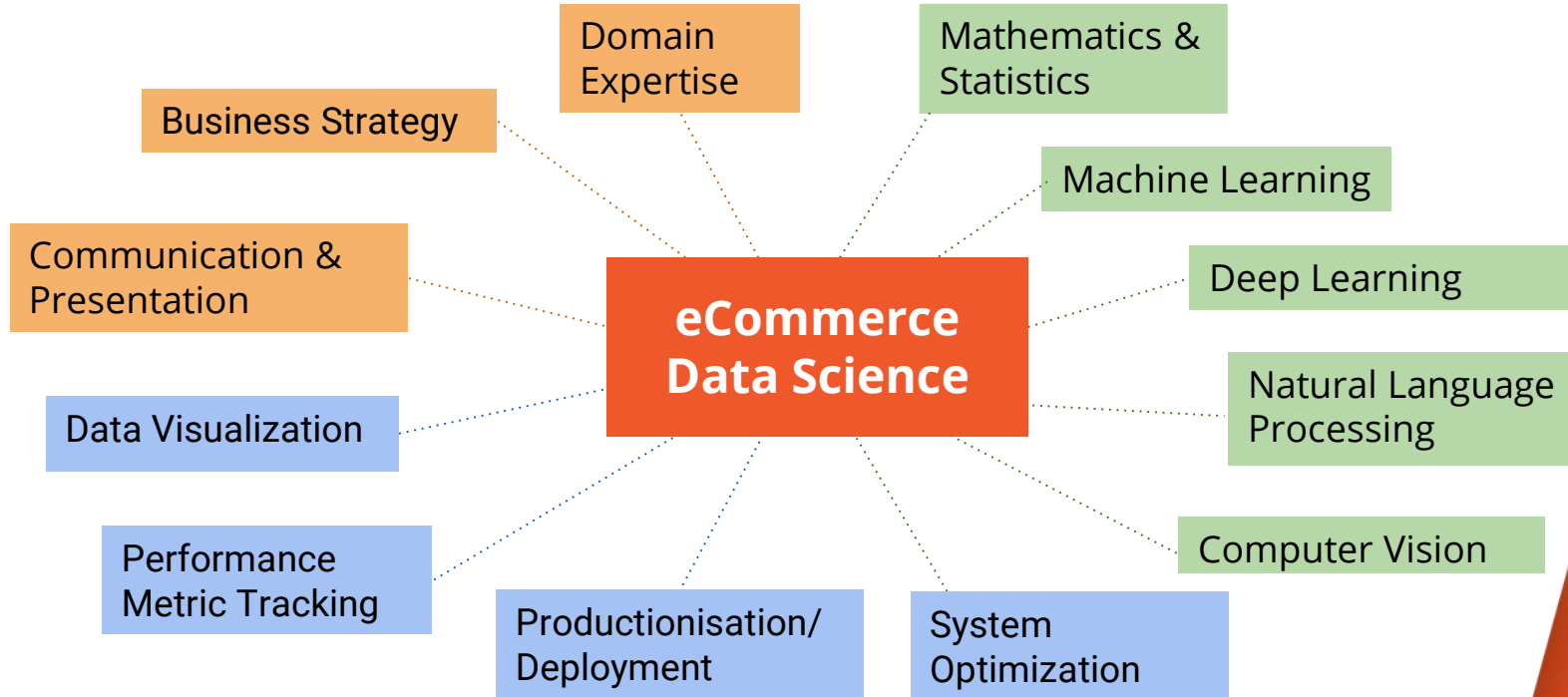
23 February 2019



What is **Data Science**?



Data Science at Shopee



Mission of Data Science at Shopee

- Drive Shopee towards **big data** driven business by using best practices of **AI technology**
- Serve our **customers better** using Data Science



Challenge that Shopee faces

One fundamental challenge is to **accurately extract relevant product information** from the large volume of products uploaded to our platform, so as to:

- Make relevant products easily discoverable through search engine and recommendation
- Improve customer experience by providing comprehensive and accurate product information



Our Problem:

Product Information Extraction in the Wild

Through this competition, we hope to inspire participants to build Data Science solutions with:

- **Computer vision**
- **Natural language processing**
- **Real world, hands-on experiences**
- **A creative mindset**



Beginner Category:

Product Category Classification

Participants are required to determine the **category** of a product given its **image and title**



The Application:

Category Recommendation

Shopee > Mobile & Gadgets > Mobile Phones & Tablets > Apple > Apple iPhone 7



www.mobilesquare.com.sg



Apple iPhone 7

\$840.00-\$1,040.00

★★★★★ 5 out of 5 (76 Ratings)

👉 Buy and earn 100 Shopee Coins ?

🚚 Shipping Fee: \$0.00 ✓

Variations:

32GB (Silver)

32GB (Gold)

128GB (Rose Gold)

128GB (E

256gb (Black)

128gb (Jet Bla



Advanced Category:

Product Information Extraction

Participants are required to detect the **brand**, **model**, and **other attributes**



The Application:

Attribute Tagging

Each product has several attributes that can provide detailed and rich information about the product



Apple iPhone X, Fully Unlocked 5.8", 64 GB - Silver

Brand: Apple

Type: Iphone X

Screen Size: 5.8"

Memory Size: 64 GB

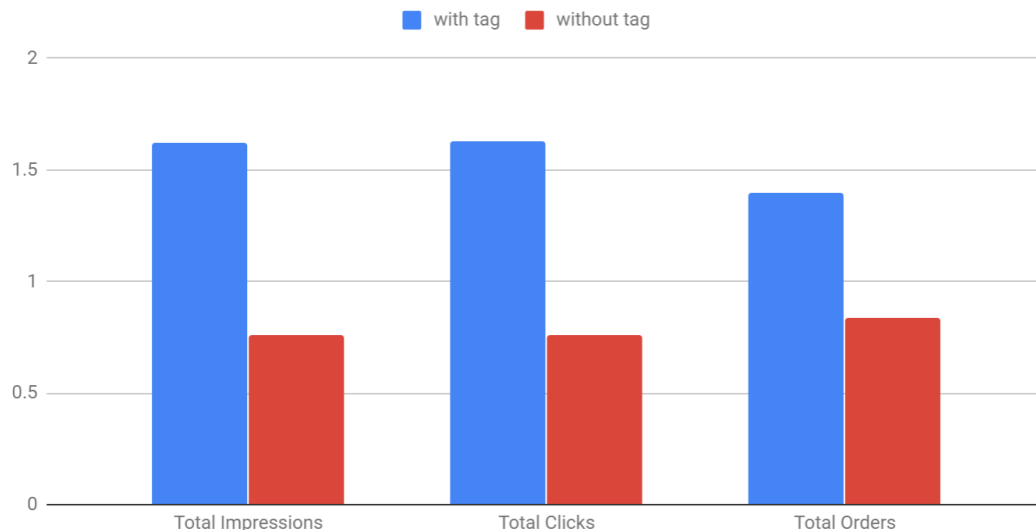
Color: Silver

Fig 1 Attributes for Iphone



Why **Attribute Tagging** is Important

Low visibility without stock keeping unit (SKU) attribute tagging results in **low conversion**



Compare Visibility of Products with and without attribute tagging



Our Data

The **product diversity** in Shopee is huge

In NDSC 2019, we select the data from three representative and prevalent product categories: **Fashion, Beauty, and Mobile**



Data Format

Training Data:

- itemid - the serial id of the item
- title - name of the item
- image_path - path for getting the product image
- category- category of the item (for beginner category)
- attribute type1, attribute type 2, ... (for advanced category)

Testing Data:

- itemid
- title
- image_path



Submission Requirements

All submissions must follow the same directory structure and result format:

- **Experiment Result**
- **Source code**
- **Supplementary Material**



How to do well in the **NDSC 2019**?

1. Deep understanding of the **e-Commerce** business
2. Good **research ability** and familiarity with **advanced scientific methods** such as Natural Language Processing, Computer Vision etc
3. Good **engineering capability** in order to convert the scientific methods into program and system



**Thank You
&
All the Best!**

