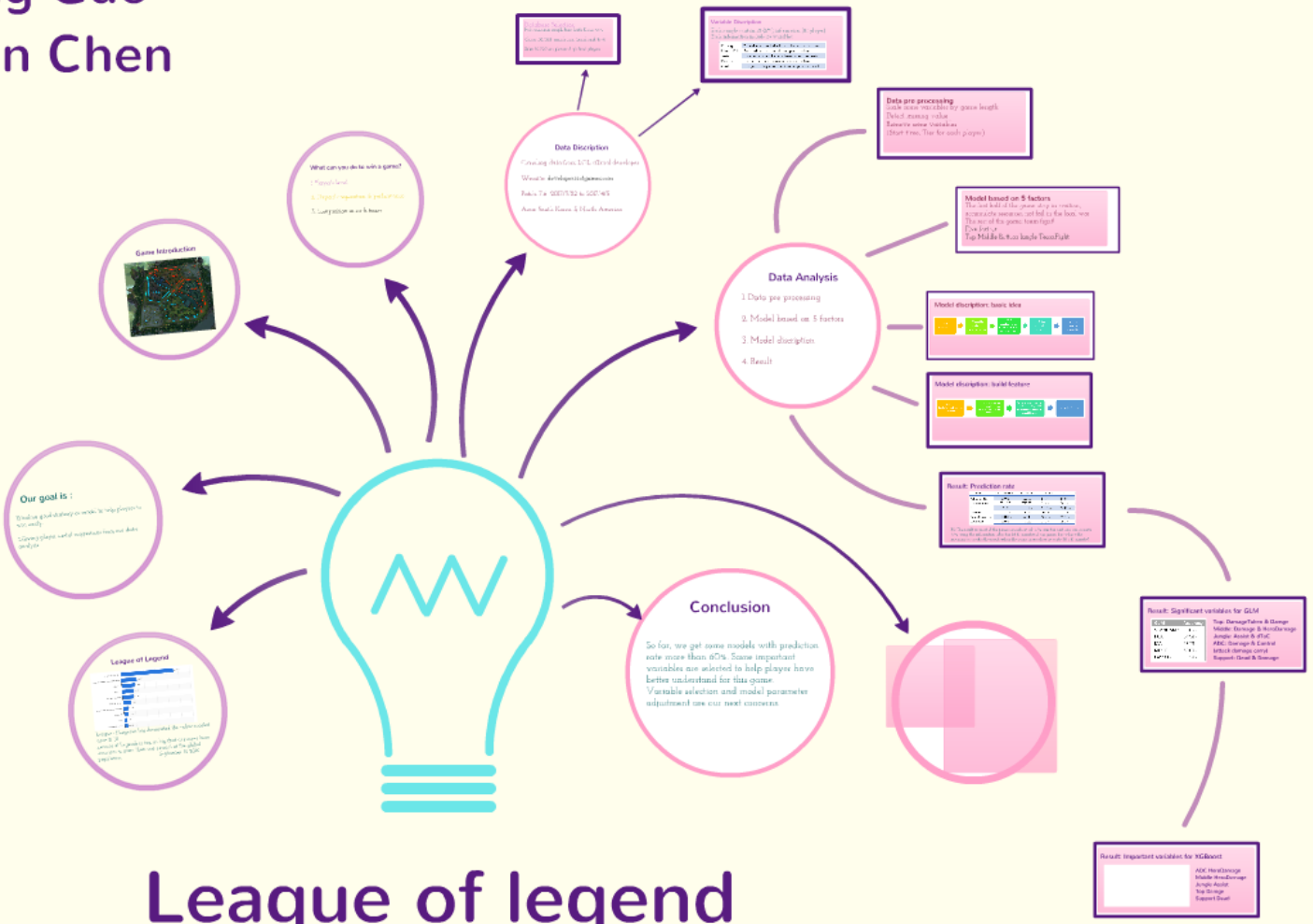
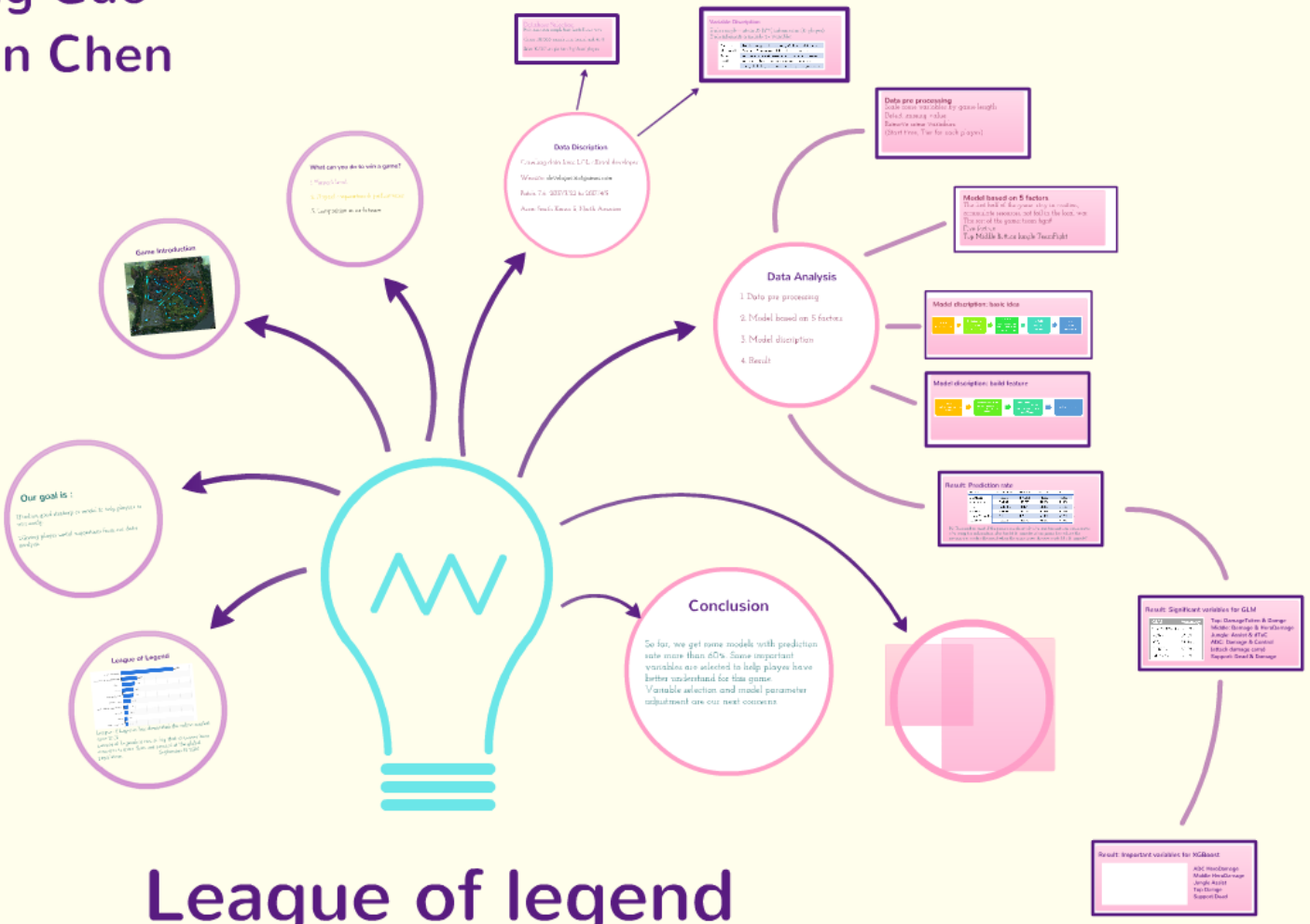
Linhai Zhang
Wen Huang
Jiacheng Xu
Zong Guo
Yilun Chen



League of legend

Thank You

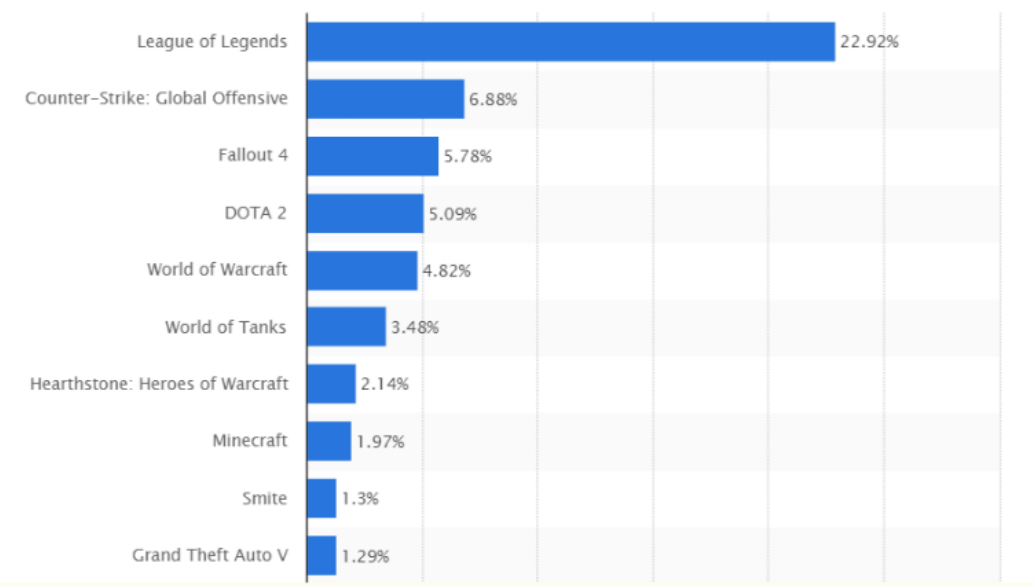
Linhai Zhang
Wen Huang
Jiacheng Xu
Zong Guo
Yilun Chen



League of legend

Thank You

League of Legend



League of Legends has dominated the online market since 2012

League of Legends is big, so big that its player base amounts to more than one percent of the global population. September 15 2016

Our goal is :

1.Finding good strategy or model to help players to win easily.

2.Giving player useful suggestions from our data analysis

Game Introduction

What can you do to win a game?

1. Player's level

2. Players' cooperation & performance

3. Composition in each team

Data Discription

Crawling data from LOL official developer

Website: developer.riotgames.com

Patch: 7.6 2017/3/22 to 2017/4/5

Area: South Korea & North America

Database Selection

Pick 1,000,000 sample from South Korea %>%

Choose 100,000 sample from formal rank %>%

Select 10,000 sample from high level players

Variable Discription

Each sample contain 10 (2*5) information (10 player)
Each information include 29 variables

Damage	Total damage includes Hero, Minion, and Structure
Minion Kill	Get gold from minion kill and push the lane
Assist	Your teammate kill enemy hero under your assist
Death	You can't obtain the resource from the lane
Gold	Use gold to buy equipment and improve yourself

Data Analysis

1. Data pre processing

2. Model based on 5 factors

3. Model discription

4. Result

Data pre processing

Scale some variables by game length

Detect missing value

Remove some variables

(Start time, Tier for each player)

Model based on 5 factors

The first half of the game: stay in position, accumulate resources, not fail in the local war

The rest of the game: team fight!

Five factors:

Top Middle Bottom Jungle TeamFight

Model discription: basic idea

Model discription: build feature

Model discription: build feature

Result: Prediction rate

Method	Crossvalidation	Boosting	All variables	Class
GLM(logit)	TRUE	FALSE	51.1%	58.8%
Randomforest	FALSE	TRUE	51.9%	56.4%
C5	FALSE	TRUE	51.4%	56.0%
RIPPER	FALSE	TRUE	50.2%	55.5%
Neural Network	TRUE	FALSE	50.4%	53.2%
XGBoost	TRUE	TRUE	51.2%	59.3%

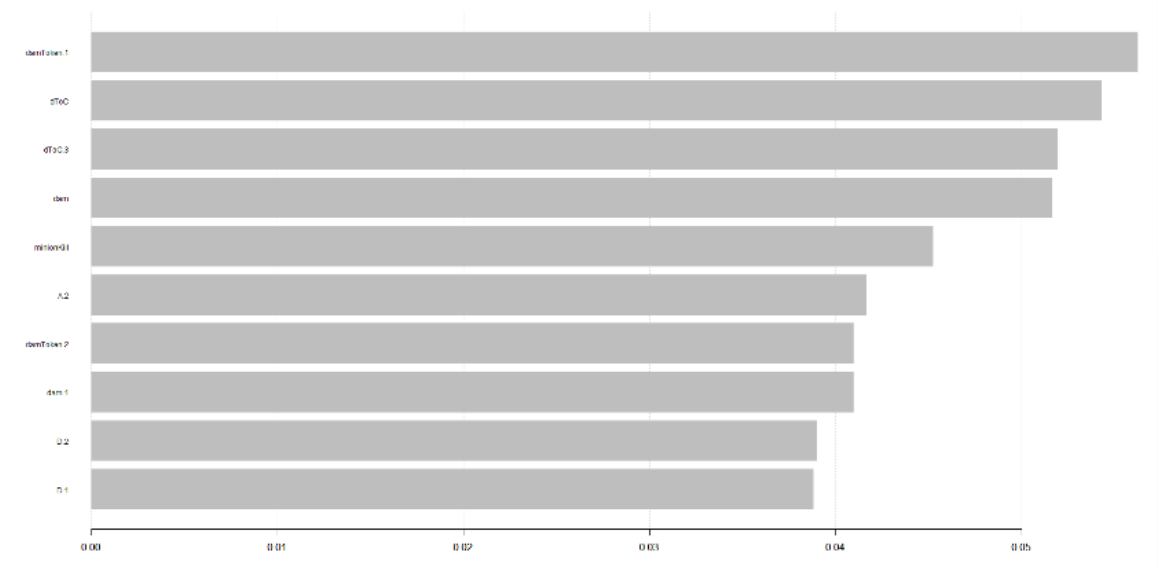
PS: The result in most of the papers are about 52~57% and the best one can achieve 67% using the information after the 1st 10 minutes of the game, but what's the meaning to predict the result when the game have already started for 10 minute?

Result: Significant variables for GLM

GLM	Accuracy
STANDARD	58.8%
PCA	54.3%
ICA	53.7%
RIDGE	63.0%
LASSO	63.3%

Top: DamageToken & Damge
Middle: Damage & HeroDamage
Jungle: Assist & dToC
ADC: Damage & Control
(attack damage carry)
Support: Dead & Damage

Result: Important variables for XGBoost



ADC HeroDamage
Middle HeroDamage
Jungle Assist
Top Damge
Support Dead

Conclusion

So far, we get some models with prediction rate more than 60%. Some important variables are selected to help player have better understand for this game.
Variable selection and model parameter adjustment are our next concerns.

Thank You