LIFELENZ
Live. Life. Better.

Final Presentation

LifeLenz App Usage (Store Type) 1

December 6th 2020

Xiuyuan Cai   Quehe Feng   Lingfeng Wu   Haibo Yu   Yifan Zhao

Xiuyuan Cai

Quehe Feng

Lingfeng Wu

Haibo Yu

Yifan Zhao

LIFELENZ

Live. Life. Better.

CONTENTS

LIFELENZ
Live. Life. Better.

Business Objectives

Company Overview:

Positive Circularity Inc

Positive Circularity invented a an automated AI platform for planning, scheduling, workforce and human capital management called as LIFELENZ.

Business Objectives:

- Finding insights on user generated data
- Offer recommendations based on findings

LIFELENZ
Live. Life. Better.

Business Objectives

Importance of Data Analysis:

Current goal for LIFELENZ is to grow its users group both in The U.S and Australia.

User satisfaction are most important for attracting more users. In order to provide better product and services to the users, finding potential improvements for users can help company entering a data spinning wheel which is also a positive circulation.

LIFELENZ

Live. Life. Better.

Data Overview

Two datasets: 1 use for general analysis, 2 use to identify store type

1

App Data

- 1,273,968 rows
- 41 columns

2

AOS Data

- 40 rows
- 4 columns

LIFELENZ
Live. Life. Better.

Data Source: Google Analytics Data
Provided by Sponsor: LIFELENZ

Data Features

Structured Features:

- **Location data**

  (country, region, metro, city)
- **Time data**

  (TimeStamp,TimeOnPage)
- **Categorical data**

  (EventAction,EventCategory)
- **Numerical data**
  (Session count, Hit order)

LIFELENZ
Live. Life. Better.

Data Features

Unstructured Features:

- **Miss values**

  (eventvalue, dimensions)
- **Unclear meaning**

  (dimension1,2,3,4)
- **Useless features**

  (browserSize, language)
- **Hidden features**
  (pagePath, secondPagepath)

LIFELENZ
Live. Life. Better.

Data Preprocessing

Data Cleaning:

- **We need use Business ID to identify store types**

Extract Business ID from pagepath column

- **We need label each Business ID as AOS or not**

Map AOS store into dataset AOS =1 or 0

- **We need clean some data for modeling**

Drop messy row and column for modeling

LIFELENZ
Live. Life. Better.

Data Preprocessing

Feature Engineering:

Features:

1. Create binary features: AOS:

   AOS=1 indicates that store is using automated optimized scheduling

2. More binary features on Event Actions:

   shift-action=1 indicates that user is using shift function on the platform

   Same as other features such as *timepunch-details*

   We created 47 variables as for 47 different event actions

LIFELENZ
Live. Life. Better.

Model Consideration

❖ Clustering

We believe that the implementation of AOS would affect the behavior pattern of the employee and manager in each stores, for example less shift and drop action.

What we want to do is let the cluster model divide the sample into two groups by behavior patterns (action list), which represent AOS store actions and None-AOS store actions.

K-means Clustering

Hierarchical Clustering

Action List → **AOS Store + None-AOS Store**

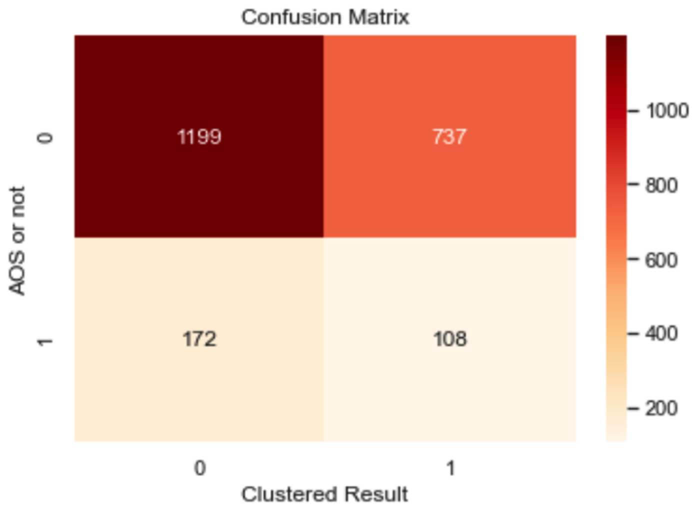
K = 2

LIFELENZ
Live. Life. Better.

Model Consideration

What we expect to find is that among the two groups of classification, one group contains most of the behaviors labeled with AOS, while the other group contains most of the behaviors labeled with non-AOS.

We plot the confusion matrix to visualize the clustering result.



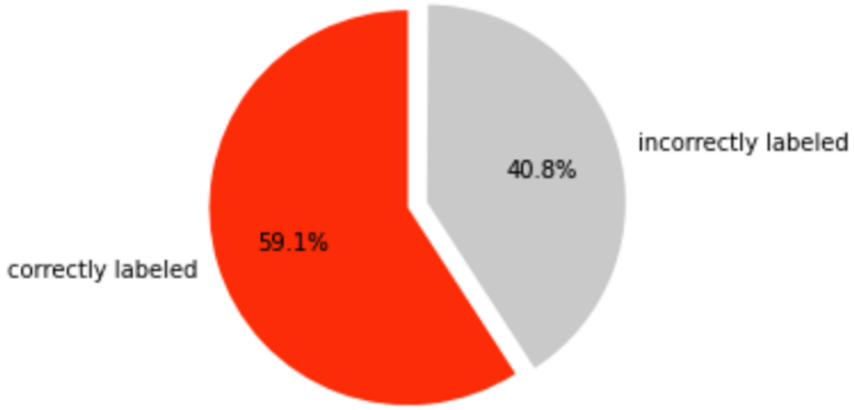
Confusion Matrix of
Clustering using K-means (k=2)

LIFELENZ
Live. Life. Better.

Model Consideration

However, the result is not ideal, and the proportion of clustering results consistent with the AOS tags is not large.

This may be due to the matrix sparseness caused by the lack of action records



Accuracy of
Clustering using K-means (k=2)

LIFELENZ
Live. Life. Better.

Model Consideration

❖ Feature Importance

We use classifiers to predict whether the action record came from an AOS Store with behavior pattern, but predicting is not our goal here and no high accuracy is needed.

```
tree = DecisionTreeClassifier(max_depth=None, min_samples_split=2, random_state=0)
tree_scores = cross_val_score(tree, x_dropped, if_aos, cv=10)
print("tree accuracy: %0.2f (+/- %0.2f)" % (tree_scores.mean(), tree_scores.std() * 2))
```

```
tree accuracy: 0.67 (+/- 0.09)
```

```
rf = RandomForestClassifier(n_estimators=10, max_depth=None,
     min_samples_split=2, random_state=0)
rf_scores = cross_val_score(rf, x_dropped, if_aos,cv=10)
print("rf accuracy: %0.2f (+/- %0.2f)" % (rf_scores.mean(), rf_scores.std() * 2))
```

```
rf accuracy: 0.73 (+/- 0.08)
```

```
et = ExtraTreesClassifier(n_estimators=10, max_depth=None,
     min_samples_split=2, random_state=0)
et_scores = cross_val_score(et, x_dropped, if_aos, cv=10)
print("et accuracy: %0.2f (+/- %0.2f)" % (et_scores.mean(), et_scores.std() * 2))
```

```
et accuracy: 0.74 (+/- 0.07)
```

```
from xgboost import XGBClassifier
num_round = 100
xgbc =XGBClassifier(max_depth=2, learning_rate=1, n_estimators=num_round, # need more weak classifiers to get fi
         objective='binary:logistic')
```

```
xgbc_scores = cross_val_score(xgbc, x_dropped.values, if_aos, cv=10)
```

```
print("xgbc accuracy: %0.2f (+/- %0.2f)" % (xgbc_scores.mean(), xgbc_scores.std() * 2))
```

```
xgbc accuracy: 0.74 (+/- 0.08)
```

LIFELENZ
Live. Life. Better.

Model Consideration

Similar to clustering, we give the task of discovering internal patterns to machine learning, just export the classifier Feature Importance to learn which features (actions) contribute most to the result (AOS or not).

randomforest

```
rf_fi.sort_values(ascending=False)
```

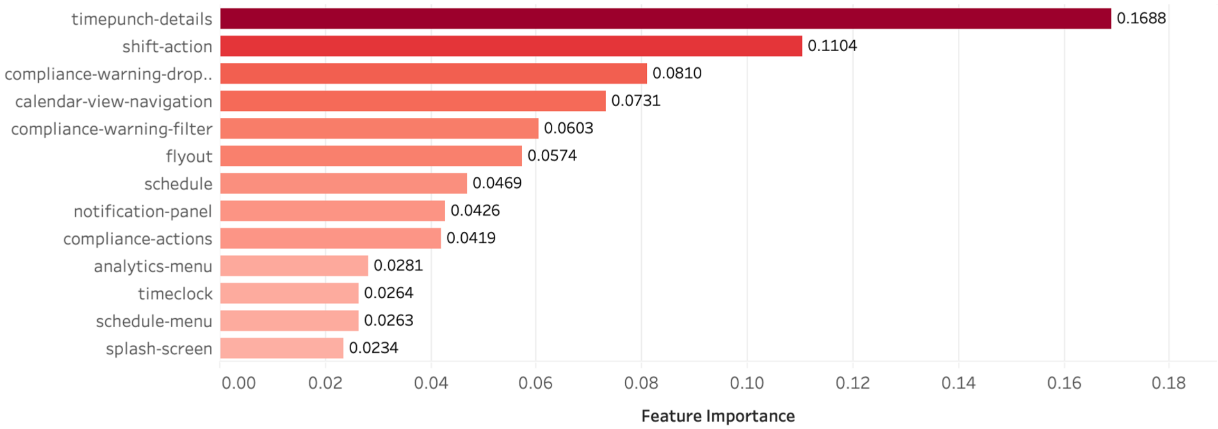
```
timepunch-details                    0.168841
shift-action                         0.110372
compliance-warning-dropdown          0.081008
calendar-view-navigation             0.073091
compliance-warning-filter            0.060347
flyout                               0.057367
schedule                             0.046941
notification-panel                   0.042645
compliance-actions                   0.041894
analytics-menu                       0.028087
timeclock                            0.026370
schedule-menu                        0.026258
splash-screen                        0.023382
manager-approval - shift-drop        0.019708
auto-create-breaks                   0.019265
```

LIFELENZ

Live. Life. Better.

Model Consideration

Now we can tell if AOS is implemented in the store, what difference in actions does the system brings to the employees



XGBoost Feature Importance

LIFELENZ
Live. Life. Better.

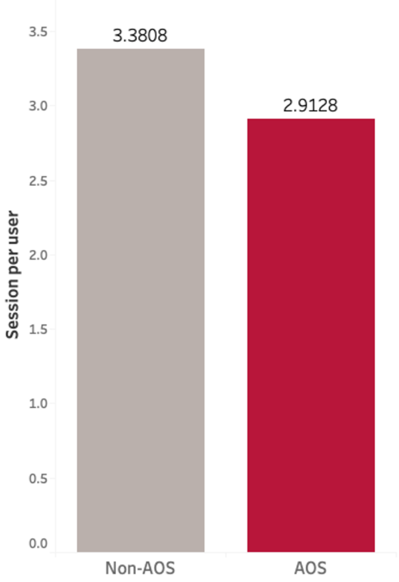
Findings and Insights

❖ Difference of Session Count

When company running AOS, per user's session count decreased from 3.4 to 2.9, which means users can spend less time for scheduling.

It significantly reduce the time users spend on scheduling and associated administrative tasks.

What Is the Difference of Session Count Between AOS and Non-AOS Store?



LIFELENZ
Live. Life. Better.

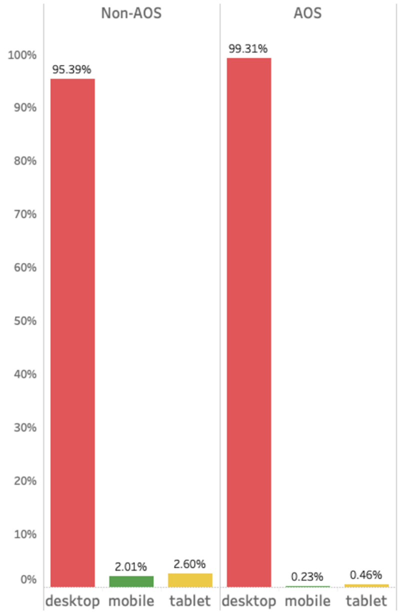
Findings and Insights

❖ Difference of Devices

More users operate at the desktop, which is over 99%, after running AOS, and less than 1% users operate at the mobile and tablet.

Therefore, LIFELENZ can put more effort and aim at the software at desktop to improve the user experience.

What Is the Difference of Devices Between AOS and Non-AOS Store?



LIFELENZ
Live. Life. Better.

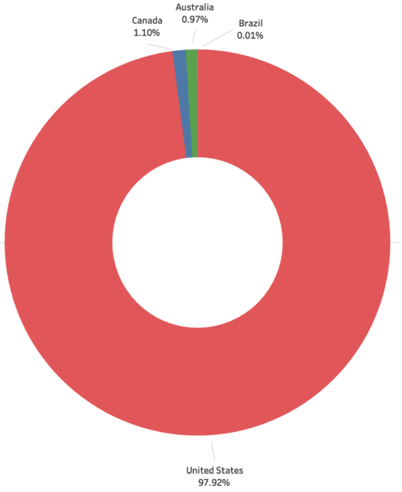
Findings and Insights

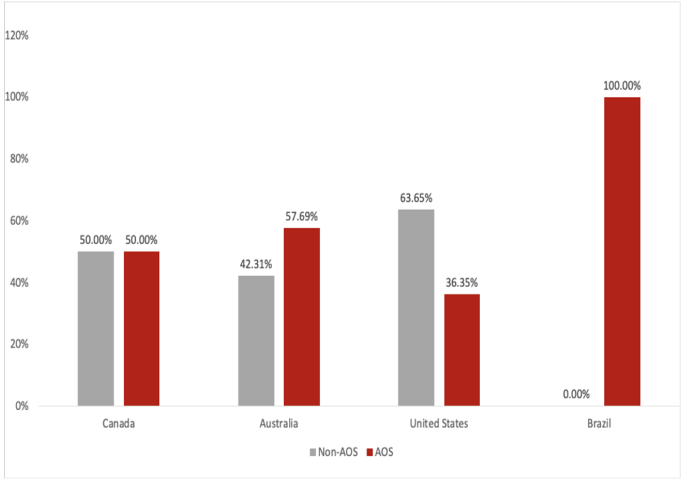
❖ **Difference of Software and AOS Usage**

Although the United States account for 98% of software usage, only 36.35% users use AOS, which is the lowest in four countries.

However, AOS is implemented well in other countries, especially in Brazil.x

What Is the Difference of Software and AOS Usage By Country?





LIFELENZ
Live. Life. Better.

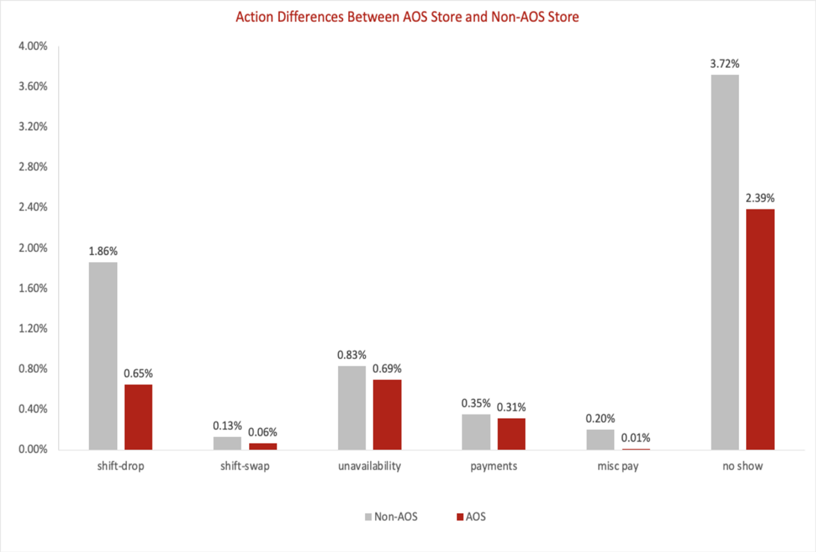
Findings and Insights

❖ Difference of Action

Although total shift actions remain the same for AOS and non-AOS stores, some actions have much differences, especially miscellaneous pay action and shift-drop rate, which reduced 96% and 65%.

Besides, no-shows rate reduced 36%, swap rate reduced 53%, unavailability rate reduced 16% and payments action rate reduced 12%.

	Shift Actions
Non-AOS	37.61%
AOS	37.41%



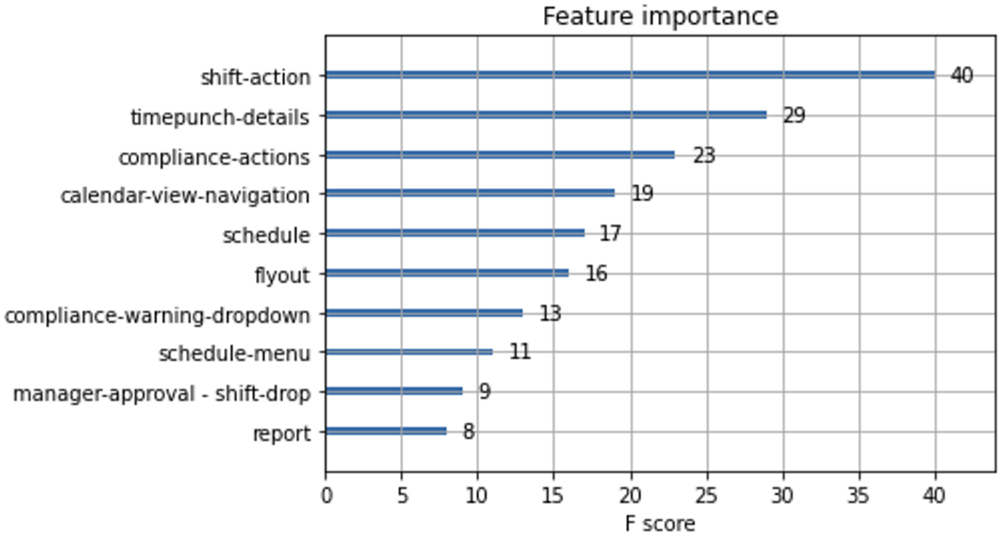
LIFELENZ
Live. Life. Better.

Findings and Insights

❖ Feature importance

After modeling, we got the feature importance of each action, and it tell us that shift action is the most important feature, and next are timepunch details and compliance action.

Then we will analyze these actions based on the feature importance.



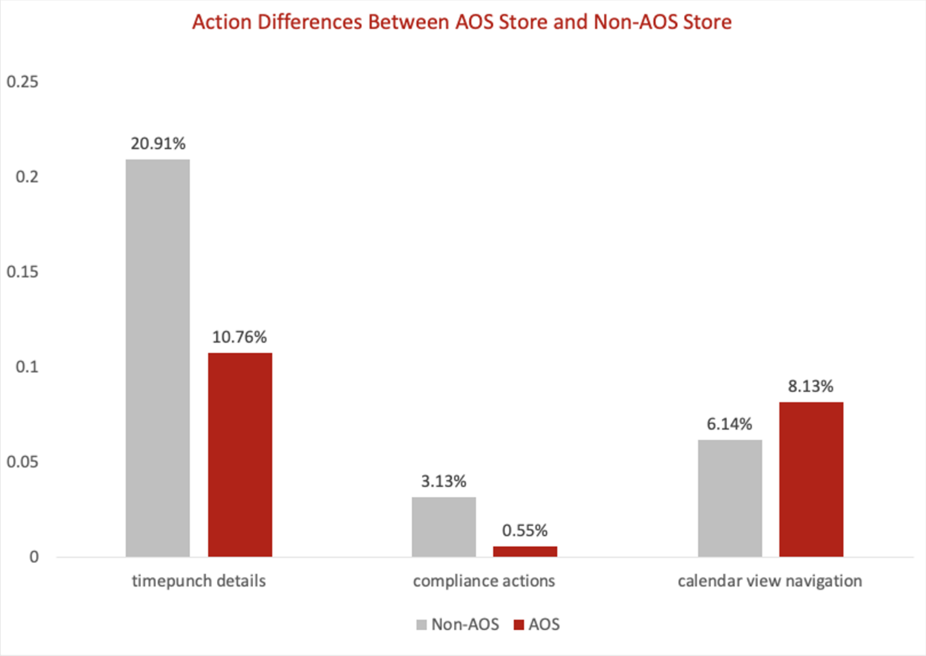
LIFELENZ
Live. Life. Better.

Findings and Insights

❖ **Difference of Action (detailed)**

We extract three actions after the model result, and we find that timepunch details decreased 49% and compliance actions decreased 82%. For calendar view navigation, AOS stores increased 32%.

We guess that increase maybe users want to check the schedule because of the automatic scheduling.



LIFELENZ
Live. Life. Better.

Recommendations and Opportunities

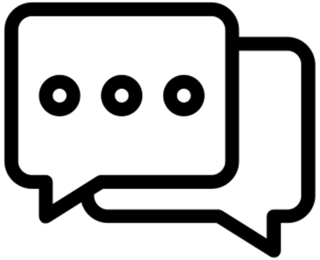
- **Adding functions in APP:**

Employee aspect: set **reminders** for each shift

lower no-shows rate

Both aspects: add a **chat box**

improve communication efficiency



LIFELENZ

Live. Life. Better.

Recommendations and Opportunities

- ❑ **Adding functions in APP:**

Manager aspect: in “Scheduling” interface, add an **suggested number of employees**

⇩

ensure sufficient labor
and lower unavailability rate

Use **A/B Test** when adding functions:

1. **Choose a possible idea after basic research**
2. **Apply ideas in product and collect signal information**
3. **Analyze information and make final decisions**

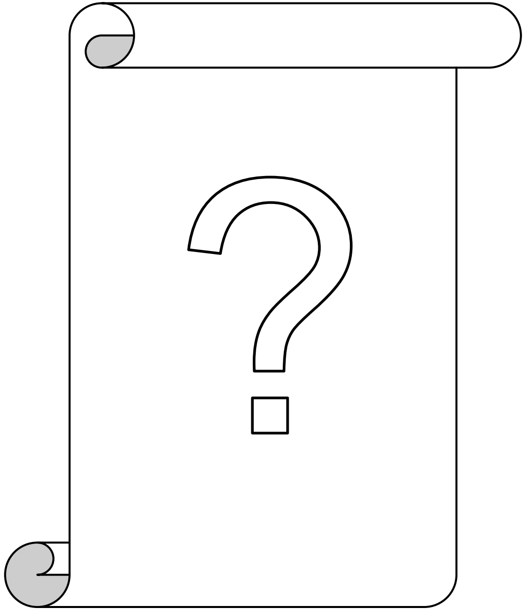
LIFELENZ
Live. Life. Better.

Recommendations and Opportunities

❑ **Improving profit and amount of users by following methods:**

1. Provide a **three-month free trial** for enterprises not using AOS, especially those in **US**.

1. Send **simple questionnaires** to both managers and employees in app, finding out the key value of LIFELENZ.



LIFELENZ
Live. Life. Better.

Recommendations and Opportunities

❑ **Future investigation:**

1. The cause of “compliance actions”

1. Why AOS stores have more “calendar view navigation” than non-AOS stores?

1. Reasons of managers reviewing timepunch details

LIFELENZ
Live. Life. Better.

Content Expansion

New Considerations for COVID-19

- Provide customers product learning to adapt to new ways of working

- Securely track employees health-related interactions

  Add new functions like ‘Tracing Route’

LIFELENZ
Live. Life. Better.

Content Expansion

New Considerations for COVID-19

❖ Assist customers to protect their employees with safer workplaces

Add new functions like ‘Healthy Self-check’

How do you feel today compared to yesterday ?

Not Well

Very Well

LIFELENZ

Live. Life. Better.

Challenges and Workarounds

Lack of professional knowledge in scheduling optimization management and practical experience

- ➢ Search the Internet for answers for definitions and relevant concepts
- ➢ Discuss controversial issues with teammates actively

Limited understanding of the exactly meaning of several columns

- ➢ Focus on particular values and find patterns to make guesses
- ➢ The demo that shows how the scheduling platform work from the perspective of manager and employee respectively

LIFELENZ
Live. Life. Better.

Challenges and Workarounds

Cannot determine which variables and features have a significant influence

- ➢ **Before establishing model: conduct exploratory data analysis to summarize the main characteristics**
- ➢ **During the modeling process: uncertain variables added to analyze the impact and decide whether to add to the model**

LIFELENZ
Live. Life. Better.

Challenges and Workarounds

Hesitating about which method is appropriate to get the optimal model

- ➢ **Selected multiple methods that fit the data characteristics to build up models** (K-means, Hierarchical, Logistic regression, Random forest, ect.)

- ➢ **Pick the one with the highest value**

LIFELENZ
Live. Life. Better.

Challenges and Workarounds

Hard to determine which methods for model evaluation

- The most commonly evaluation methods are Accuracy and ROC-AUC curve for the binary classification
- More intuitive and easily understand to use accuracy and ROC Curve to measure the quality of models
- Both Accuracy and ROC curve are used to measure models and take the model with the highest value

LIFELENZ
Live. Life. Better.

THAN
K
YOU!

LifeLenz App Usage (Store Type) 1

LIFELENZ

Live. Life. Better.

Final Presentation Video Link:

https://drive.google.com/file/d/1J06llEtvEbDWW_XHjy1BVUoYC0QrLxu6/view?usp=sharing