

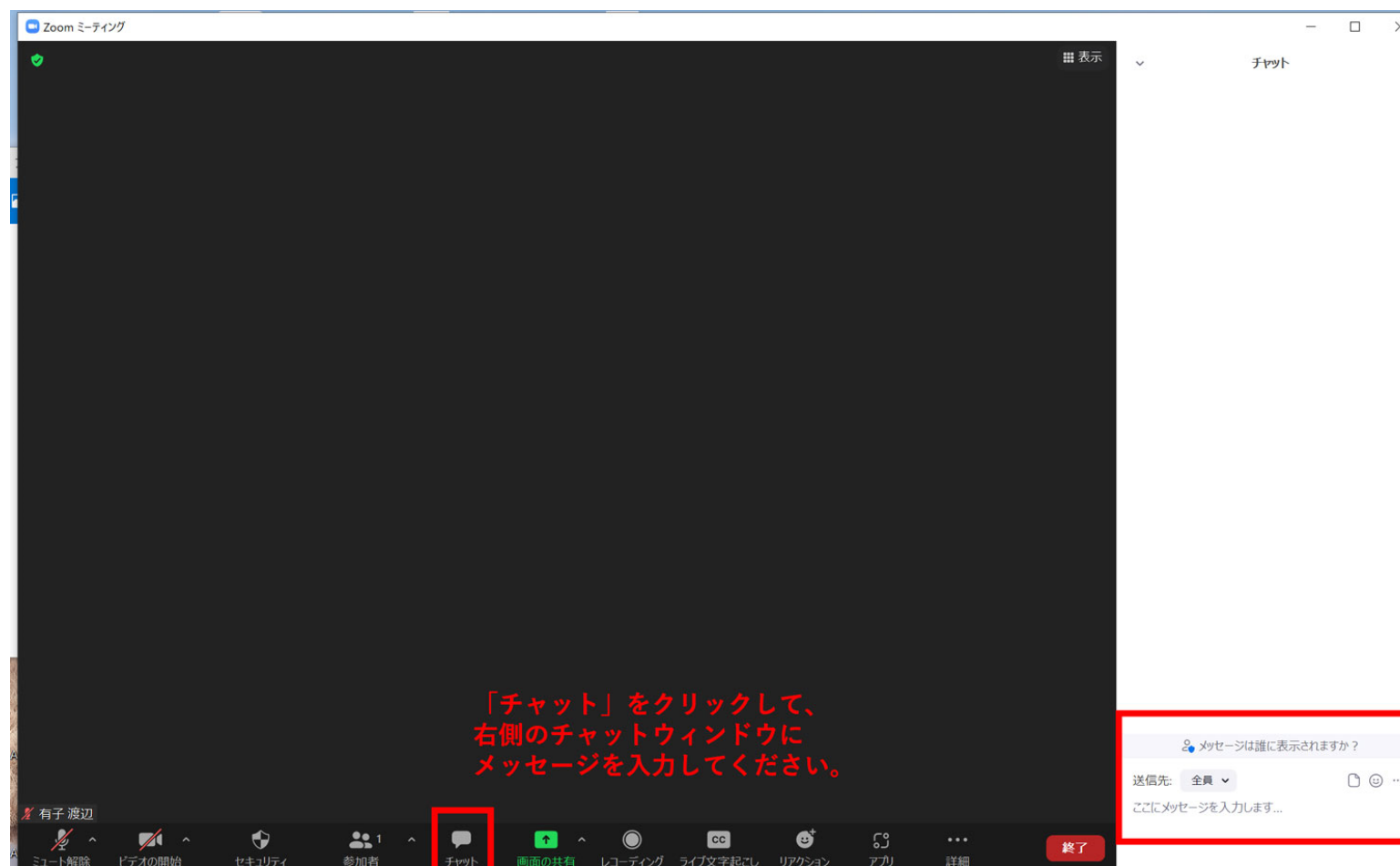
# UCI Bank

## 定期預金申し込み顧客の分類デモ

### ★RapidMinerGO★



# ご質問はチャットからお気軽にどうぞ



# アンケートへのご協力をお願いいたします。

Zoomから退出すると、アンケートフォームのURLが表示されます。「続行」をクリックして簡単なアンケートにご協力をいただければ幸いです。



# はじめに

- デモの内容  
RapidMinerGOで定期預金申し込み顧客を分類するデモ手順
- 分析要求  
定期預金申し込み確度の高い顧客を分類すること
- 必要なデータセット（ 事前にインストールをして、デスクトップ上に保存 ）
  - 学習データ : 01\_bank\_training.xlsx
  - テストデータ : 02\_bank\_test.xlsx

# BANK データのフォーマット

エクセルデータをあらかじめデスクトップ上にダウンロードをお願いいたします。詳細はこちらのページから  
<https://digital-labor.nissho-ele.co.jp/course-rapidminer/rapidminer-ai-begginer>

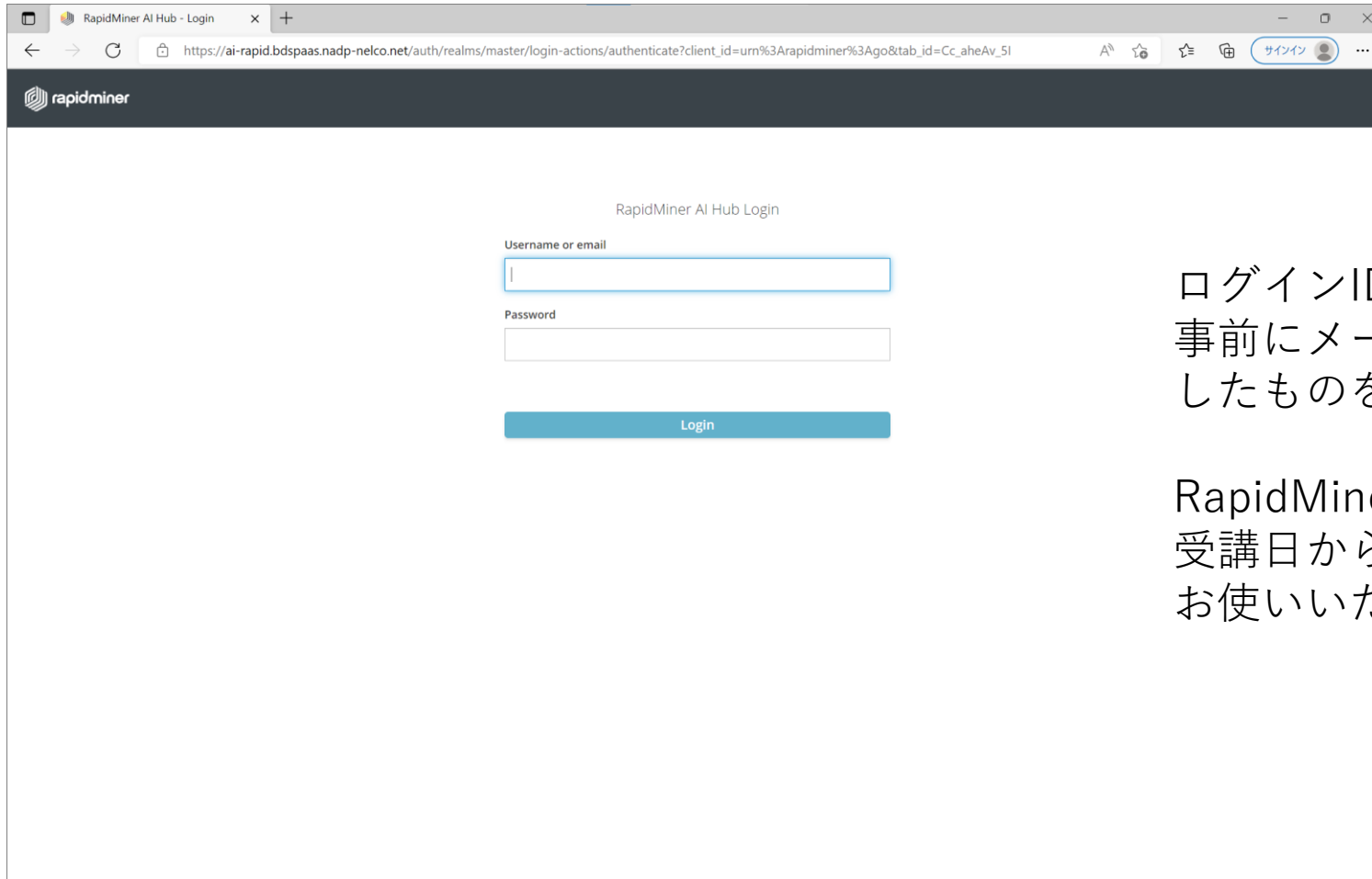
No	項目名	意味	
①	年齢	年齢	顧客情報
②	職種	職種	
③	結婚歴	結婚歴	
④	学歴	学歴	
⑤	債務不履行	債務不履行の有無	
⑥	平均残高	年間平均残高	
⑦	住宅ローン	住宅ローンの有無	
⑧	個人ローン	個人ローンの有無	
⑨	連絡手段	連絡手段	
⑩	最終接触日	最後に接触した日付（日）	キャンペーン情報
⑪	最終接触月	最後に接触した日付（月）	
⑫	最終接触j時間	接触した時間（秒）	
⑬	今回接触回数	今回のキャンペーンでの接触回数	
⑭	前回経過日数	前回のキャンペーンでの接触後の経過日数	
⑮	前回接触回数	今回のキャンペーン以前の接触回数	
⑯	前回預金申込	前回のキャンペーンの成功有無	目的変数
⑰	預金申込	預金申し込みの有無	

# GOデモの流れ



# RapidMinerGOへログイン

URL <https://ai-rapid.bdspaas.nadp-nelco.net/am/>



The screenshot shows a web browser window with the title 'RapidMiner AI Hub - Login'. The address bar displays the URL: [https://ai-rapid.bdspaas.nadp-nelco.net/auth/realms/master/login-actions/authenticate?client\\_id=urn%3Arapidminer%3Ago&tab\\_id=Cc\\_aheAv\\_5l](https://ai-rapid.bdspaas.nadp-nelco.net/auth/realms/master/login-actions/authenticate?client_id=urn%3Arapidminer%3Ago&tab_id=Cc_aheAv_5l). The page content includes the RapidMiner logo and the text 'RapidMiner AI Hub Login'. Below this, there is a login form with two input fields: 'Username or email' and 'Password'. A blue 'Login' button is positioned below the password field.

ログインIDとパスワードは  
事前にメールにてお知らせ  
したものをお使いください。

RapidMinerGOはセミナー  
受講日から2週間、ご自由に  
お使いいただけます。

★初回ログイン時にID・パスワードを入れると、下記の承認画面がでますので「I ACCEPT」ボタンを押してお進みください。

RapidMiner Go End User License Agreement

\*\*\* IMPORTANT \*\*\*

PLEASE READ CAREFULLY BEFORE YOU DOWNLOAD OR USE THE SOFTWARE

This document (the "Agreement") is a legal agreement between RapidMiner, Inc. ("RapidMiner") and you (the "Licensee"). The software that you are downloading and/or using (the "Software") is the exclusive property of RapidMiner or its licensors and is protected by United States and International Intellectual Property Laws. The Software is copyrighted and licensed (not sold). RapidMiner is only willing to license the Software subject to the terms and conditions of this Agreement, and any use of the Software outside of the scope of such terms and conditions is prohibited.

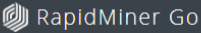


By clicking the <Accept> button or checkbox at the end of this document or by downloading, installing, copying, executing or otherwise using the Software, you acknowledge that you have read this Agreement, understand it and agree to be bound by its terms and conditions. If you are not willing to be bound by the terms of this Agreement, click the <Reject> button or

I ACCEPT

I REJECT





こちらの画面に切り替わります。では、始めて行きましょう！


RapidMiner Go 

## Welcome to RapidMiner Go, user01!


RapidMiner Go is designed to help you build predictive models from your data – fast and simple. It helps you choose the best models and apply them to new data to make predictions.




Add Data



Apply models to  
new datasets



Manage deployed  
model

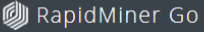




Manage analyses

### Recent Activity


You do not have any recent activity.

# Add Data をクリック

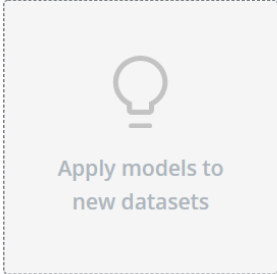
RapidMiner Go 

## Welcome to RapidMiner Go, user01!

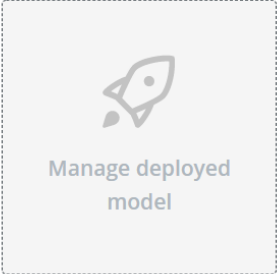
RapidMiner Go is designed to help you build predictive models from your data – fast and simple. It helps you choose the best models and apply them to new data to make predictions.



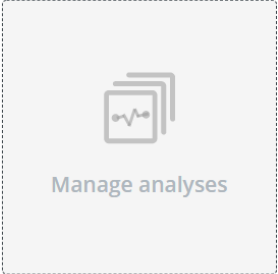
Add Data



Apply models to new datasets



Manage deployed model



Manage analyses

### Recent Activity

You do not have any recent activity.

# 01\_bank\_training.xlsxをドラッグ&ドロップ

**RapidMiner Go**

**Add your data!**

**Data set requirements:**

- File size:** up to unlimited rows and 500 columns and at most 2.00 GB
- File types:** [Read more on requirements](#)

**Add Data** **Use Sample Dataset**

Your data is safe. We care about your privacy and the integrity of your data, and we will delete the data within 28 days, upon request. [Read more](#)

**Or drag and drop your data set here**

Your dataset should look like below:  
Delimiter: "comma, semicolon, tab, or space"

	A	B	C	D
Place column names in the first row (header)				

Sheet 1 Add your data on the first sheet only in Excel

**Or choose one of your previous Dataset:**


Search



**File Explorer (AI-Seminer):**


- 01\_bank\_training
- 02\_bank\_test


2 個の項目

# データのプレビュー & Build a new predict modelをクリック


 RapidMiner Go


 **Review your data!**



Find patterns in your data



Build a new predictive model



BETA  
Build a new time series model

Preview your data [01\\_bank\\_training.xlsx](#) (400 rows, 17 columns)

	年齢	職種	結婚歴	学歴	債務不履行	平均残高	住宅ローン	個人ローン	連絡手段	最終接触日	最終接触月	最終接
	Numerical	Categorical	Categorical	Categorical	Categorical	Numerical	Categorical	Categorical	Categorical	Numerical	Categorical	Nu
1	56	事務職	既婚	中等教育	なし	45	なし	なし	?	5	5月	
2	54	事務職	既婚	高等教育	なし	184	なし	なし	?	5	5月	
3	44	技能職	既婚	?	なし	146	あり	なし	?	15	5月	
4	44	技術職	既婚	高等教育	なし	4580	あり	なし	?	15	5月	
5	30	技術職	未婚	高等教育	なし	880	あり	あり	?	19	5月	
6	59	サービス職	既婚	中等教育	なし	-22	あり	なし	?	26	5月	
7	48	技術職	既婚	中等教育	あり	0	なし	なし	?	30	5月	
8	35	サービス職	既婚	中等教育	なし	152	あり	なし	?	2	6月	
9	44	管理職	離婚	高等教育	あり	-411	あり	なし	?	2	6月	
10	24	技能職	既婚	中等教育	なし	-389	あり	あり	?	4	6月	

上部プルダウン：予測目的とする項目「預金申込」を設定  
下部プルダウン：預金申込者をターゲットとするため、「あり(申し込みする)」を設定し、次へ

ポイント  
・「あり」を設定  
する目的は何か、  
一緒に考えてみて  
ください



RapidMiner Go

1 Choose Column — 2 Select Inputs — 3 Select Models — 4 Inspect Results

< Choose Column to Predict 預金申込 Next

Great! You chose a column with only two categorical values.

You need to choose which value is more important to predict correctly. あり

① Define Gains and Costs

Distribution of the target column "預金申込"

Preview your data [01\\_bank\\_training.xlsx](#) (400 rows, 17 columns)

	年齢	職種	結婚歴	学歴	債務不履行	平均残高	住宅ローン	個人ローン	連絡手段	最終接触日	最終接触月	最終
	Numerical	Categorical	Categorical	Categorical	Categorical	Numerical	Categorical	Categorical	Categorical	Numerical	Categorical	
1	56	事務職	既婚	中等教育	なし	45	なし	なし	?	5	5月	
2	54	事務職	既婚	高等教育	なし	184	なし	なし	?	5	5月	
3	44	技能職	既婚	?	なし	146	あり	なし	?	15	5月	
4	44	技術職	既婚	高等教育	なし	4580	あり	なし	?	15	5月	
5	30	技術職	未婚	高等教育	なし	880	あり	あり	?	19	5月	
6	59	サービス職	既婚	中等教育	なし	-22	あり	なし	?	26	5月	
7	48	技術職	既婚	中等教育	あり	0	なし	なし	?	30	5月	
8	35	サービス職	既婚	中等教育	なし	152	あり	なし	?	2	6月	
9	44	管理職	離婚	高等教育	あり	-411	あり	なし	?	2	6月	

分析に使用する項目を取捨選択する。  
問題のある項目は自動的に非選択状態となっているので、  
初期値のまま次へ。



- Correlation、ID-ness、Missingとは？
  - それが品質にどう影響するか？
- 一緒に考えてみてください

RapidMiner Go

Choose Column — 2 Select Inputs — 3 Select Models — 4 Inspect Results

Select your features

RapidMiner Go removes the columns that have quality issues from your data. If you want to override RapidMiner Go's decision you can include any of the deselected columns by re-selecting it.

Filter columns: All (16) Questionable quality (1) Poor quality input (3) (13 of 16 columns selected) Search Columns

	Name	Quality	Correlation	ID-ness	Stability	Missing
<input type="checkbox"/>	債務不履行	Highly stable	0.04%	0.50%	98.25%	0.00%
<input type="checkbox"/>	連絡手段	Highly stable	6.50%	0.50%	90.19%	21.00%
<input type="checkbox"/>	前回預金申込	Many missing values	1.72%	0.75%	40.62%	76.00%
<input checked="" type="checkbox"/>	年齢	Low correlation	0.00%	13.75%	6.25%	0.00%
<input checked="" type="checkbox"/>	職種		0.62%	2.75%	27.32%	0.25%
<input checked="" type="checkbox"/>	結婚歴		1.21%	0.75%	57.25%	0.00%
<input checked="" type="checkbox"/>	学歴		0.02%	0.75%	47.23%	5.25%
<input checked="" type="checkbox"/>	平均残高		0.20%	87.00%	5.75%	0.00%
<input checked="" type="checkbox"/>	住宅ローン		6.27%	0.50%	52.50%	0.00%
<input checked="" type="checkbox"/>	個人ローン		2.51%	0.50%	86.00%	0.00%
<input checked="" type="checkbox"/>	最終接触日		2.79%	7.75%	6.50%	0.00%
<input checked="" type="checkbox"/>	最終接触月		1.49%	3.00%	24.75%	0.00%

- Quality
- poor quality ⇒ 警告。分析に適さない  
債務不履行：ほぼ「なし」  
連絡手段：ほぼ「携帯電話」  
前回申し込み：ほぼ入力がない(欠損値)
  - questionable quality ⇒ 注意。ビジネスドメインに照らし合わせて、取捨選択する
- 年齢：「預金申し込み」との相関関係 (correlation) が低いため、注意喚起されている。ビジネスドメイン、ビジネスの状況に照らして合わせて取捨選択する。今回は選択状態で分析実施。

# Show advanced settings をクリック

ポイント

- ・ Easily、Quick、Higher とは？



RapidMiner Go

Choose Column — Select Inputs — 3 Select Models — 4 Inspect Results

< Select your models Run Analysis

Are you looking for quick insights? Or do you want easily interpretable models? Choose your preferred approach:

**Machine Learning Algorithms**

- ☒ **Easily Interpretable** ⓘ  
Decision trees and simple linear models
- ☐ **Quick Prototyping** ⓘ  
Quick insights using generic models
- ☐ **Higher Accuracy** ⓘ  
Longer running time with bigger selection of models

**Training Data Selection**

Sampling Type: Automatic Sampling ⓘ

**Column Analysis**

- ☒ Importance of columns ⓘ
- ☒ Explain predictions ⓘ

**Data Preparation**

- ☒ Remove columns with too many values ⓘ ^

Show advanced settings

# 「Generalized Linear Model」、「Deep Learning」、「Gradient Boosted Trees」を選択し、Run Analysisをクリック

ポイント

- ・ 3つのアルゴリズムを選択した理由は？一緒に考えてみてください



上から

2番目「Generalized Linear Model」  
5番目「Deep Learning」  
8番目「Gradient Boosted Trees」  
を選択

RapidMiner Go

Choose Column — Select Inputs — **3 Select Models** — 4 Inspect Results

**Select your models** **Run Analysis**

Choose one or more of the following methods to model your data: Show advanced settings

**Machine Learning Algorithms**

- ☐ Naive Bayes
- ☒ Generalized Linear Model
- ☐ Logistic Regression
- ☐ Fast Large Margin
- ☒ Deep Learning
- ☐ Decision Tree
- ☐ Random Forest
- ☒ Gradient Boosted Trees

**Training Data Selection**

Sampling Type: Automatic Sampling

**Column Analysis**

- ☒ Importance of columns
- ☒ Explain predictions

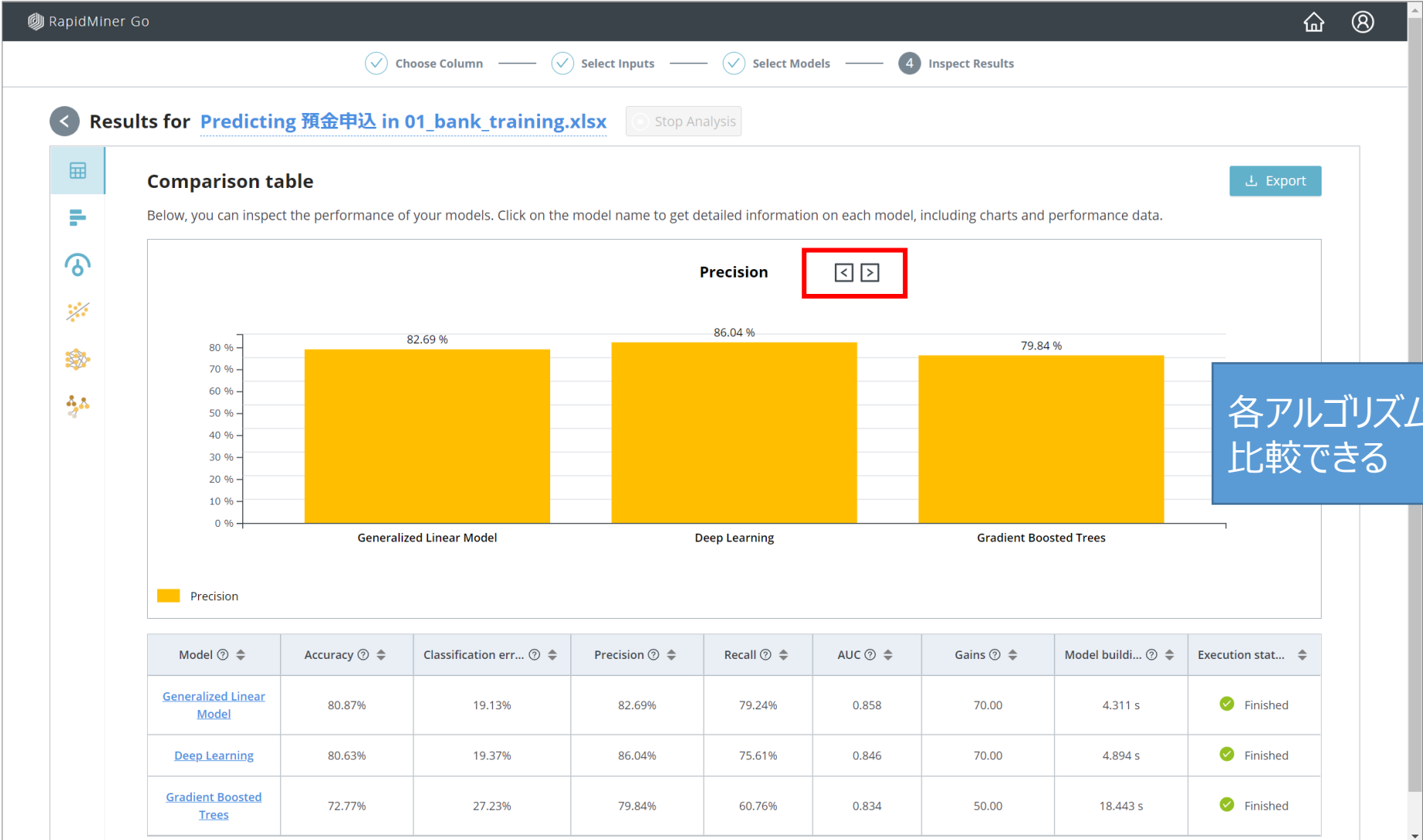
**Data Preparation**

- ☒ Remove columns with too many values

データセットの内容に応じて、推奨されるアルゴリズムが表示される。その中から代表的な3つのアルゴリズムを選択した。



選択したアルゴリズムの予測精度を比較できる  
右矢印ボタンをクリックし、「Precision = 適合率」の画面を表示  
預金申込者が正しく分類できている確率(=適合率)の最も良いモデルは「Deep Learning」

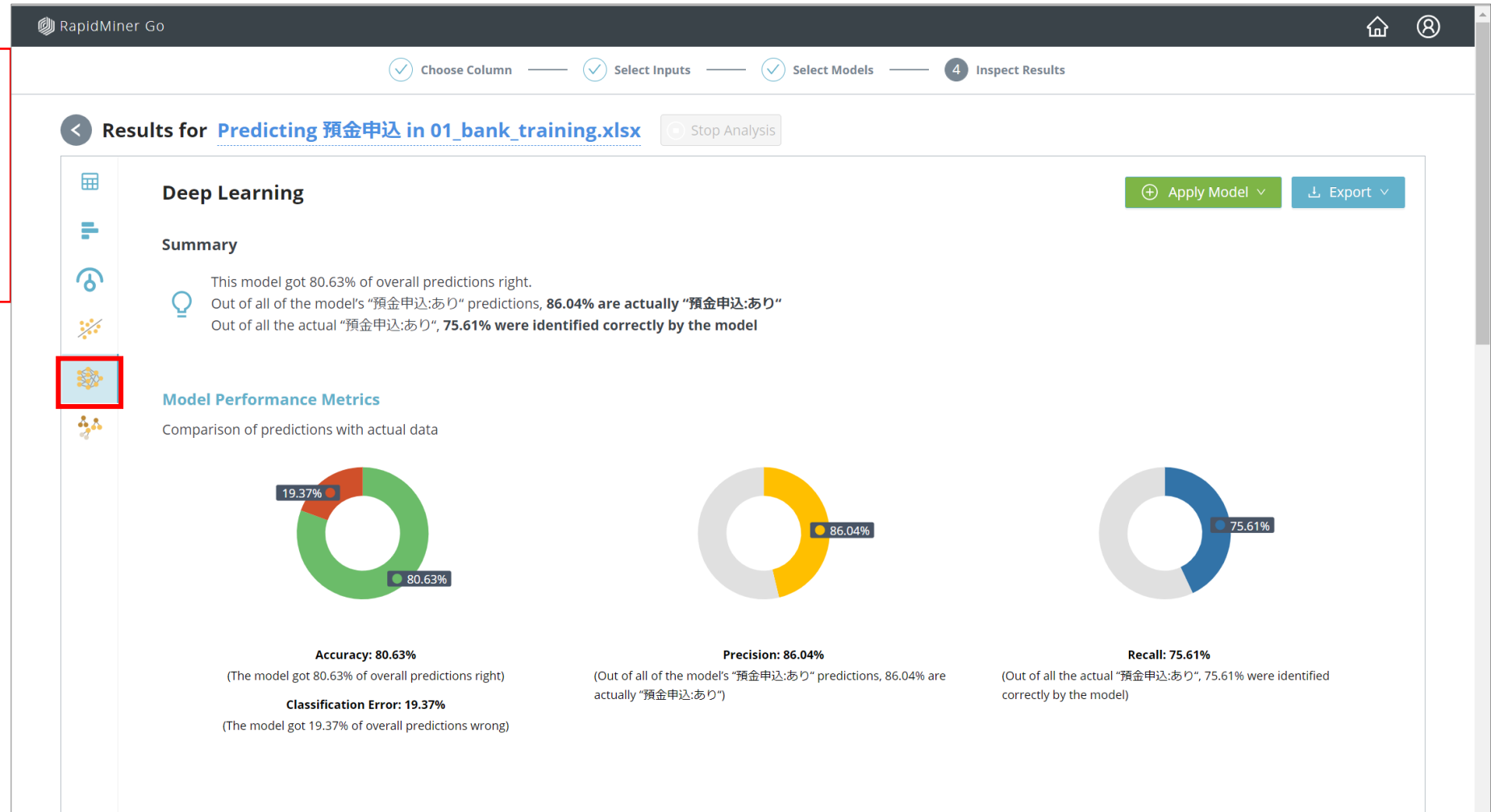


各アルゴリズムの評価指標を  
比較できる

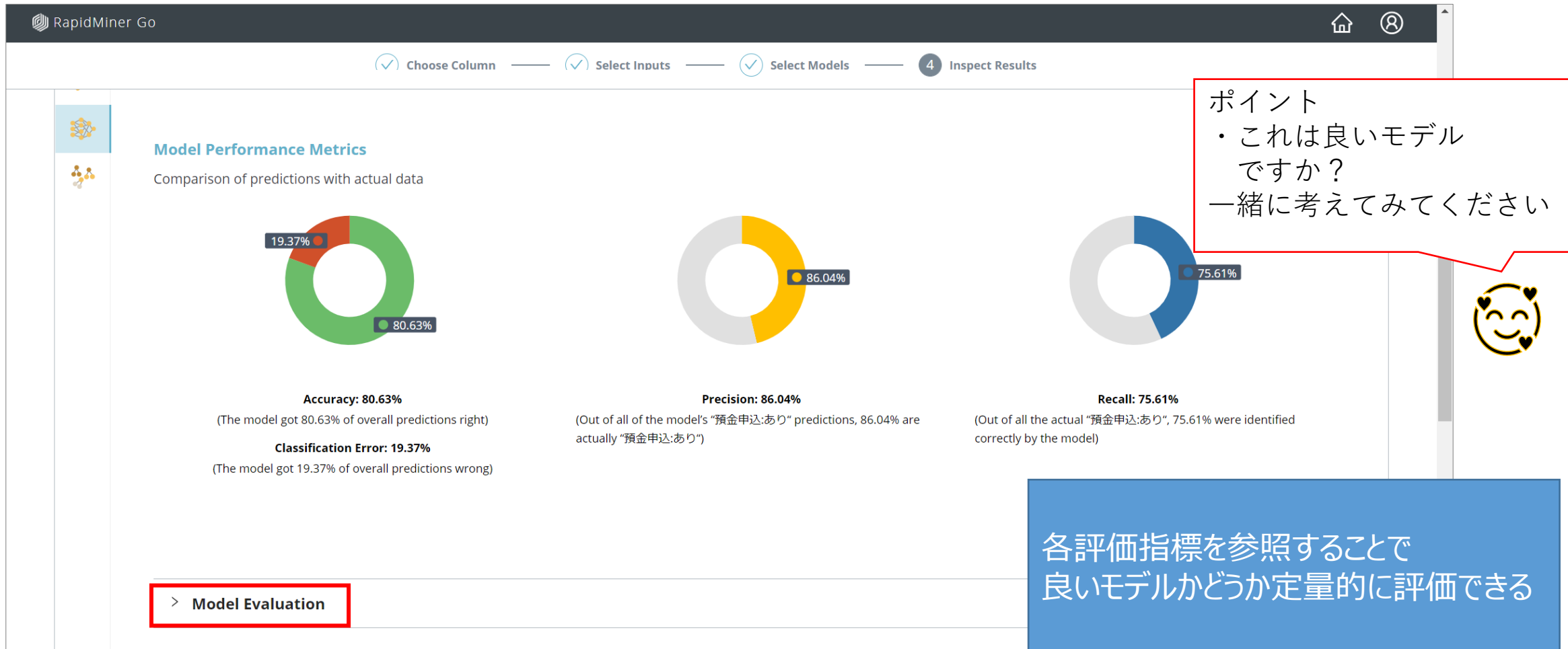
# 「Deep Learning」のボタンをクリック

## ポイント

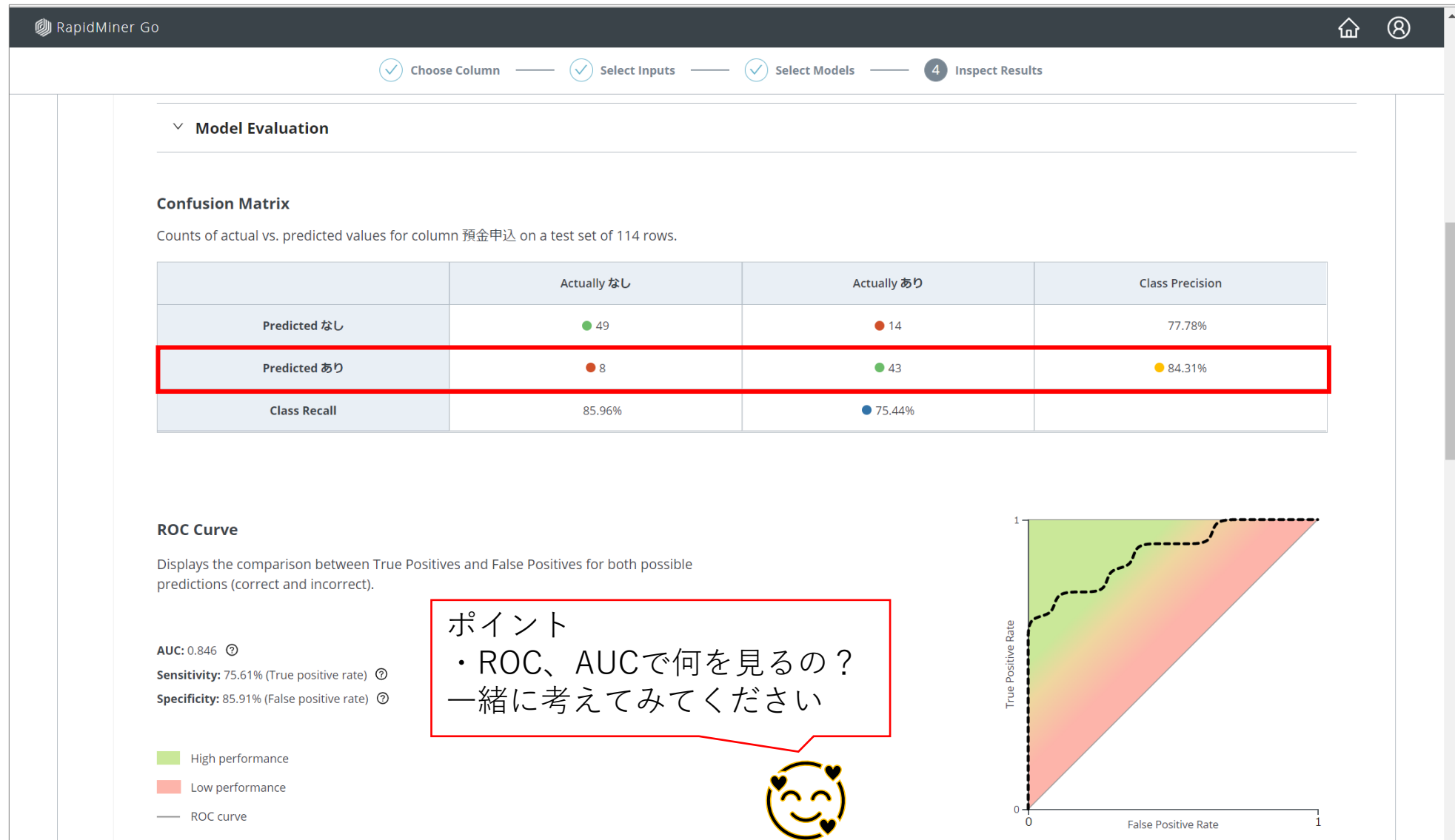
- Precision、Recallとは？
  - Accuracyでは、何故ダメなのか？
- 一緒に考えてみてください



# 「Deep Learning」の各評価指標を確認 Model Evaluationをクリック



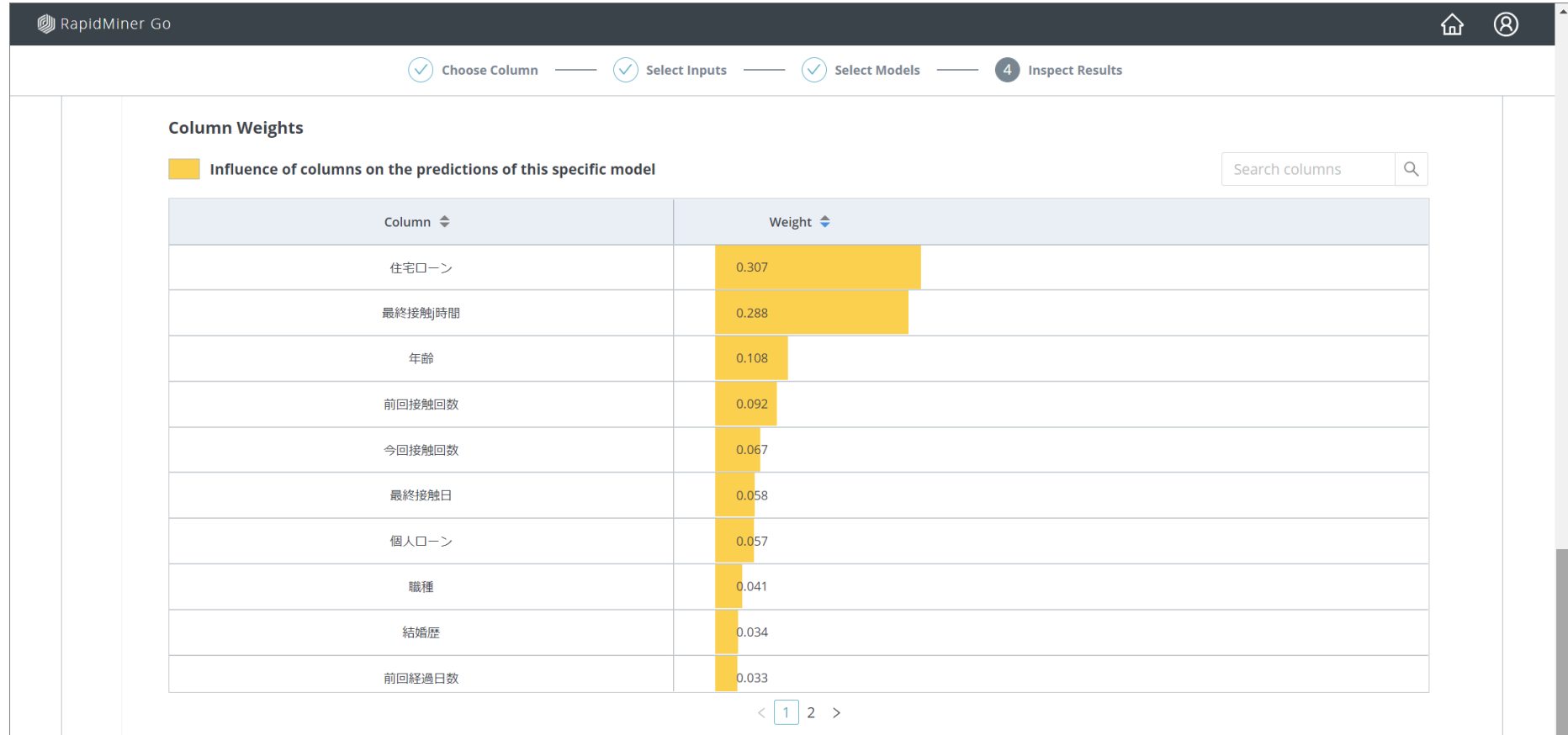
# 「Confusion Matrix」で「Precision = 適合率」の内訳を確認 預金申込者と予想した51人中43人が正しく分類できており、適合率84%であることを確認



「Column Weights」重要度が高い項目がランキング形式で確認できる  
預金申込者の分類に重要度が一番高いのは「住宅ローン、最終接触時間」  
納得感があるモデルかどうかビジネスの現場で確認する

ポイント

・Weightsで何をするの？  
一緒に考えてみてください



# 「Model Simulator」のボタンをクリック 「Choose a model」は「Deep Learning」を選択

RapidMiner Go

Choose Column — Select Inputs — Select Models — 4 Inspect Results

< Results for Predicting 預金申込 in 01\_bank\_training.xlsx Stop Analysis

### Model Simulator

Use this real-time interface to change the inputs of a model and view the output. It shows predictions, confidences and explanations for those inputs.

**Choose a model:** Deep Learning

- Generalized Linear Model
- Deep Learning**
- Gradient Boosted Trees

年齢: 42.72

学歴:

平均残高: 2068.70

前回接触回数: 0.59

職種: 管理職

最終接触月: 5月

今回接触回数: 2.42

結婚歴: 既婚

最終接触日: 16.36

個人ローン: なし

前回経過日数: 41.45

住宅ローン: なし

最終接触時間: 389.29

#### Most likely

あり

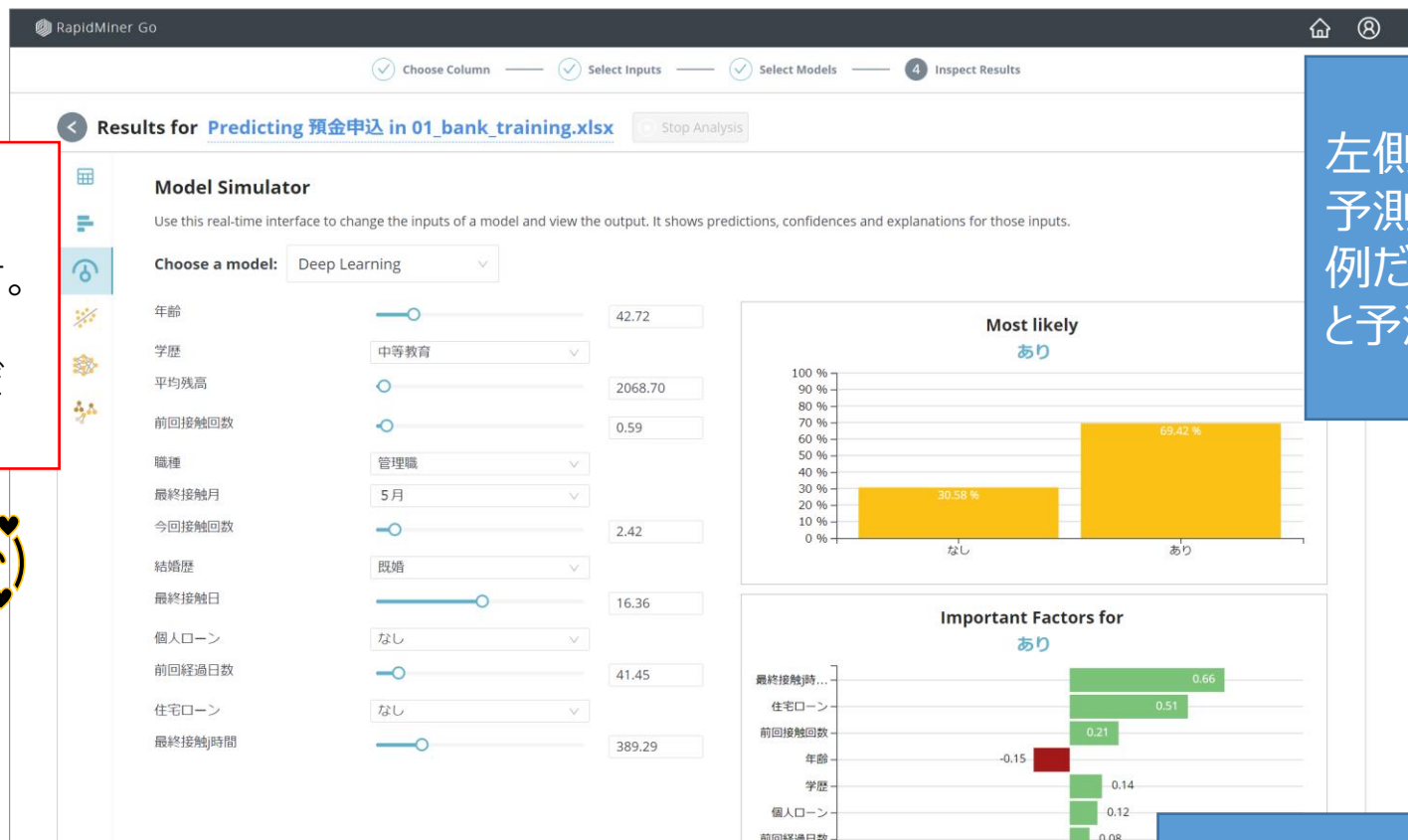
Category	Probability
なし	30.58 %
あり	69.42 %

#### Important Factors for

あり

Feature	Importance
最終接触日時...	0.66
住宅ローン	0.51
前回接触回数	0.21
年齢	-0.15
学歴	0.14
個人ローン	0.12

先ほど作成した「Deep Learning」のモデルを用いて、仮にこういう顧客属性の人であれば、どのくらいの確率で預金申込するかシミュレーションすることができます。(=what-if分析)



ポイント

- Model SimulatorのURLが共有できます。
- その目的は？一緒に考えてみてください

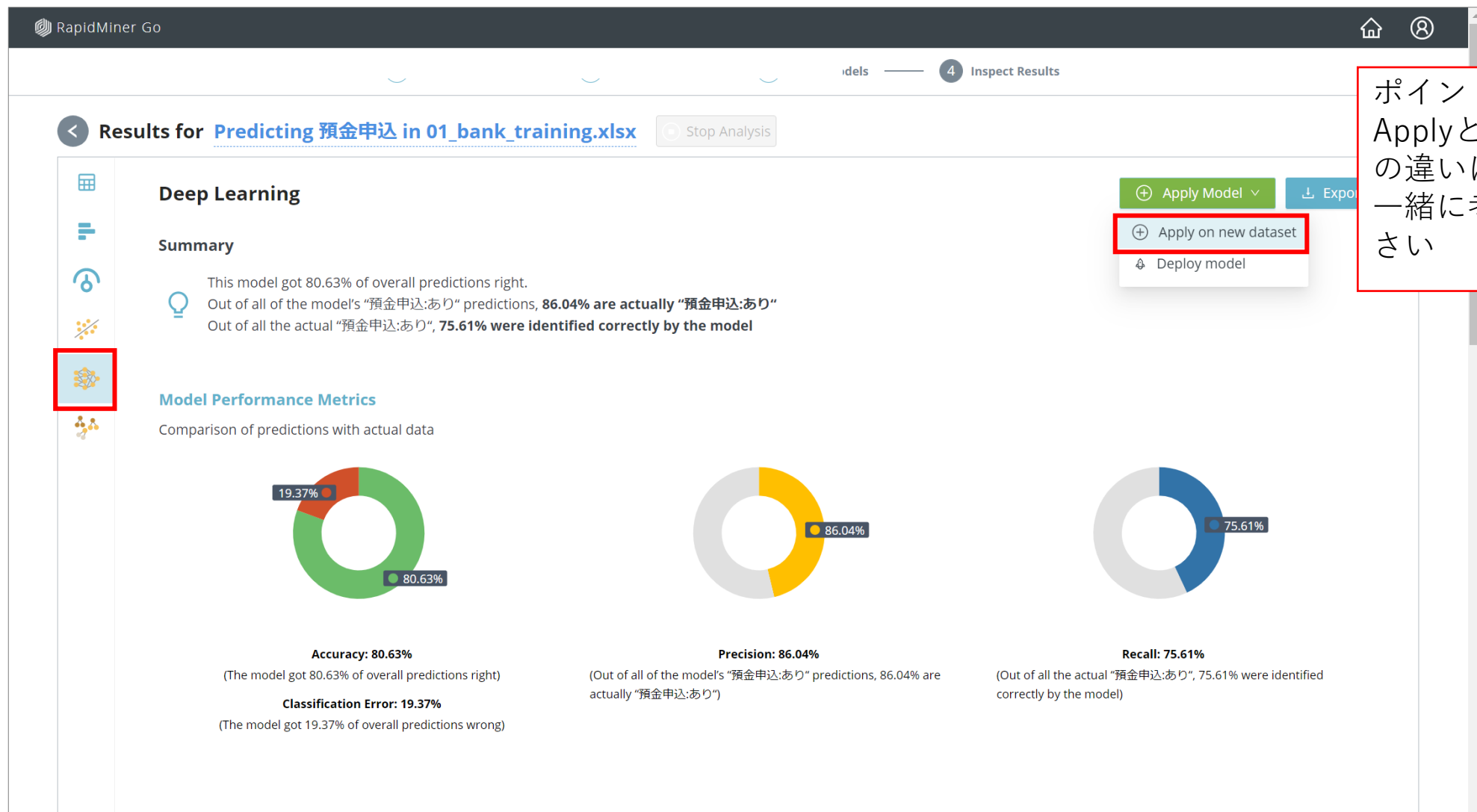


左側の画面で設定した値に応じて、予測結果をリアルタイム表示  
例だと69%の確率で預金申込すると予測されている

初期値は学習データの平均値がセットされている、1レコードがセットされた状態

予測に作用した項目が影響度が大きい順に表示  
緑：予測結果(あり)に対してプラスに作用した項目  
赤：予測結果(あり)に対してマイナスに作用した項目

「Deep Learning」をクリック  
「Apply Model」>「Apply on new dataset」をクリック



ポイント  
ApplyとDeploy Model  
の違いは？  
一緒に考えてみてくだ  
さい





## 「02\_bank\_test.xlsx」をドラッグ&ドロップ

RapidMiner Go

✓ Select Model

2 Add Data

3 Check Dataset

4 Inspect Predictions

<

Add your data!

⚙

Data set requirements:

The data you upload here should match the training data you used to build a model, with the same columns and the same data types. For example, if in the training data the "Month" column was identified as a number with values 1-12, then in the new data set "Month" should also be identified as a number, not a textual value (January, February, etc.).

[Read more on requirements](#)

+ Add Data

📄 Use Sample Dataset

ⓘ Your data is safe. We care about your privacy and the integrity of your data, and we will delete the data within 28 days, upon request.

[Read more](#)

×

Or **drag and drop** your data set here

Your dataset should look like below:  
Delimiter: "comma, semicolon, tab, or space"

	A	B	C	D
→				
Place column names in the first row (header)				

Sheet 1 ← Add your data on the first sheet only in Excel

Or choose one of your previous Dataset:

Search

🔍

📄

01\_bank\_training.xlsx

47.66 KB

2022-07-05 11:32:20

📁 | 📄 | 🗑 | AI-Seminer

—

□

×

📁

ファイル

ホーム

共有

表示

▼

?

←

→

▼

↑

📁

« デス...

»

AI-Se...

▼

📁

名前

^

☑

📄

01\_bank\_training

☑

📄

02\_bank\_test

2 個の項目

1 個の項目を選択

47.6 KB

📄

🖨

→

# 「Calculate Predictions」をクリック

ポイント

・未知のデータなのに  
正解ラベル付きなのは  
何故でしょう？

一緒に考えてみてくだ  
さい



RapidMiner Go

Select Model — Add Data — 3 Check Dataset — 4 Inspect Predictions

< Check Dataset **Calculate Predictions**

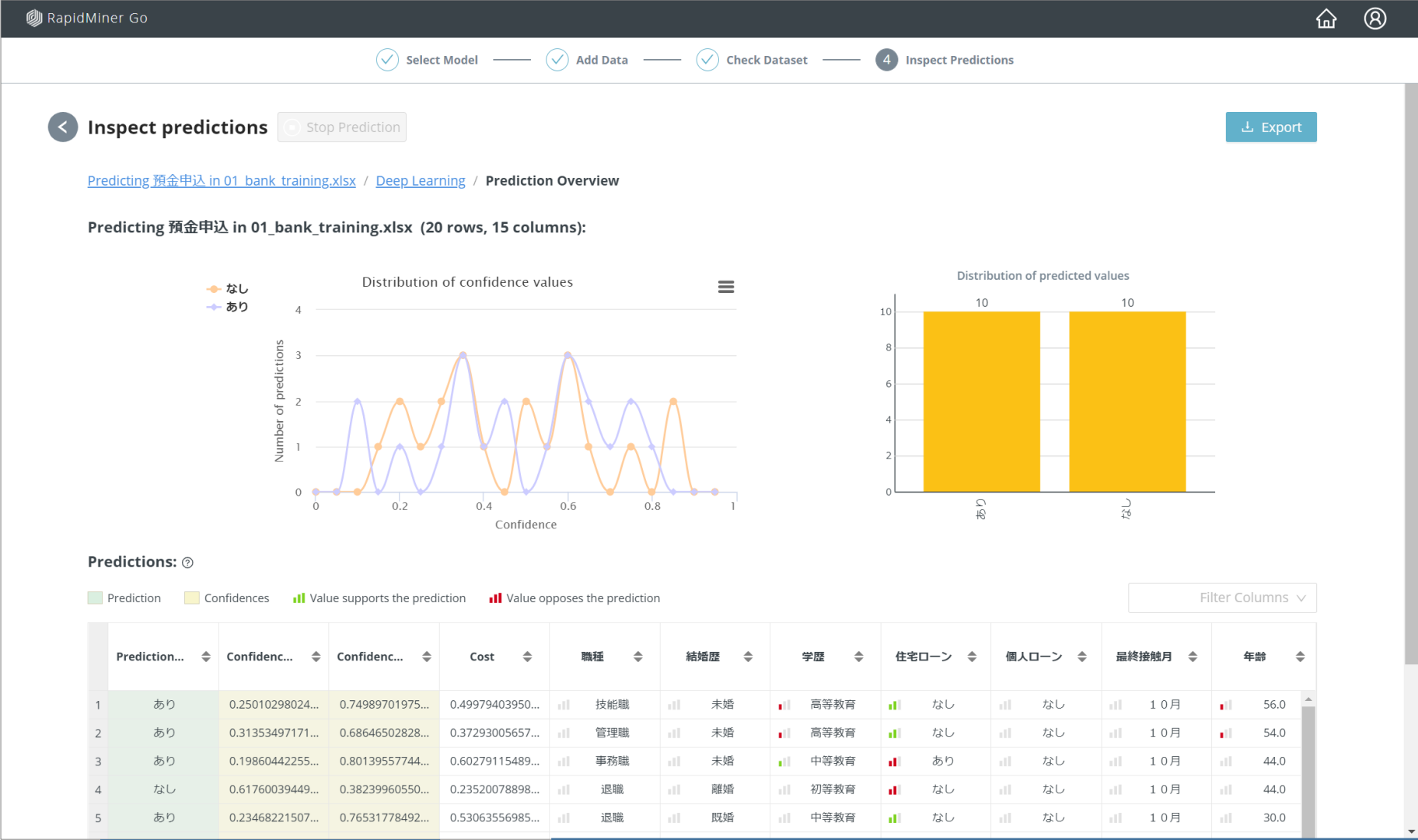
[Predicting\\_預金申込 in 01\\_bank\\_training.xlsx](#) / [Deep Learning](#) / Review Data

✔ This column is required.  
✔ This column is not required. The model ignores it when generating predictions.

Filter Columns ▾

	✔ 年齢 ▾	✔ 職種 ▾	✔ 学歴 ▾	✔ 結婚歴 ▾	✔ 債務不... ▾	✔ 平均残高 ▾	✔ 住宅口... ▾	✔ 個人口... ▾	✔ 最終接... ▾	✔ 最終接... ▾	✔ 最終接... ▾
	Numerical	Categorical	Categorical	Categorical	Categorical	Numerical	Categorical	Categorical	Numerical	Categorical	Numerical
1	56	技能職	高等教育	未婚	なし	455	なし	なし	22	1 0 月	542
2	54	管理職	高等教育	未婚	なし	2081	なし	なし	22	1 0 月	554
3	44	事務職	中等教育	未婚	なし	1693	あり	なし	22	1 0 月	608
4	44	退職	初等教育	離婚	なし	2693	なし	なし	25	1 0 月	83
5	30	退職	中等教育	既婚	なし	1495	なし	なし	25	1 0 月	389
6	59	技術職	なし	未婚	なし	11115	なし	なし	25	1 0 月	189
7	48	技術職	中等教育	既婚	なし	2779	なし	なし	25	1 0 月	324
8	35	技術職	高等教育	既婚	なし	844	あり	あり	25	1 0 月	129
9	44	学生	中等教育	未婚	なし	1234	なし	なし	26	1 0 月	280
10	24	退職	中等教育	既婚	なし	1058	なし	なし	27	1 0 月	277

# 未知のデータに対する予測結果を確認



予測  
結果

予測確率

緑の棒グラフ表示項目：予測結果に対してプラスに作用した項目  
赤の棒グラフ表示項目：予測結果に対してマイナスに作用した項目

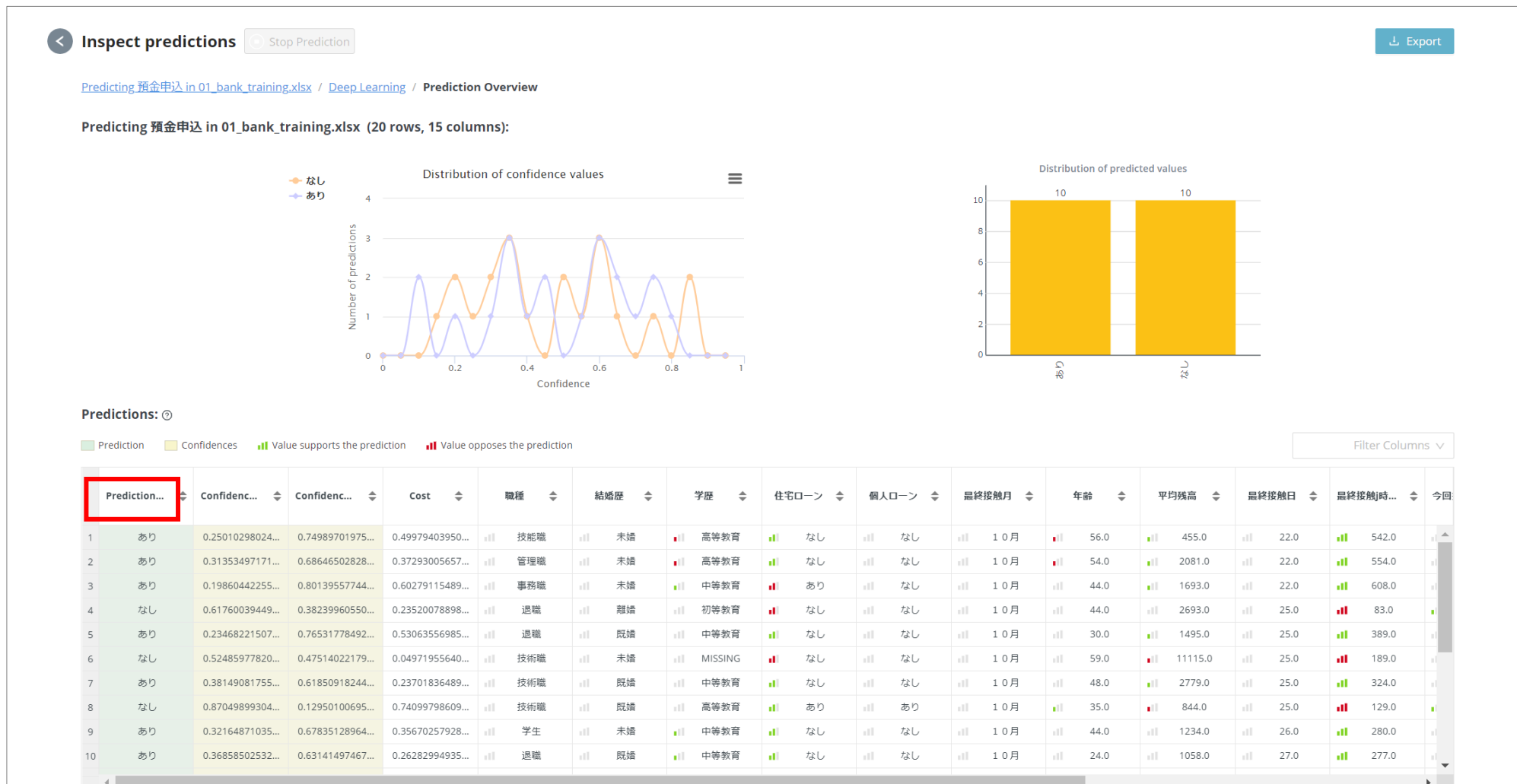
# Predictionの上矢印ボタンをクリック

預金申込者と予想した10人中6人が正しく分類できており、適合率60%であることを確認

## ポイント

- ・TPが60%ってどう？
- ・FPが40%ってどう？

などなど、結果について一緒に考えてみてください





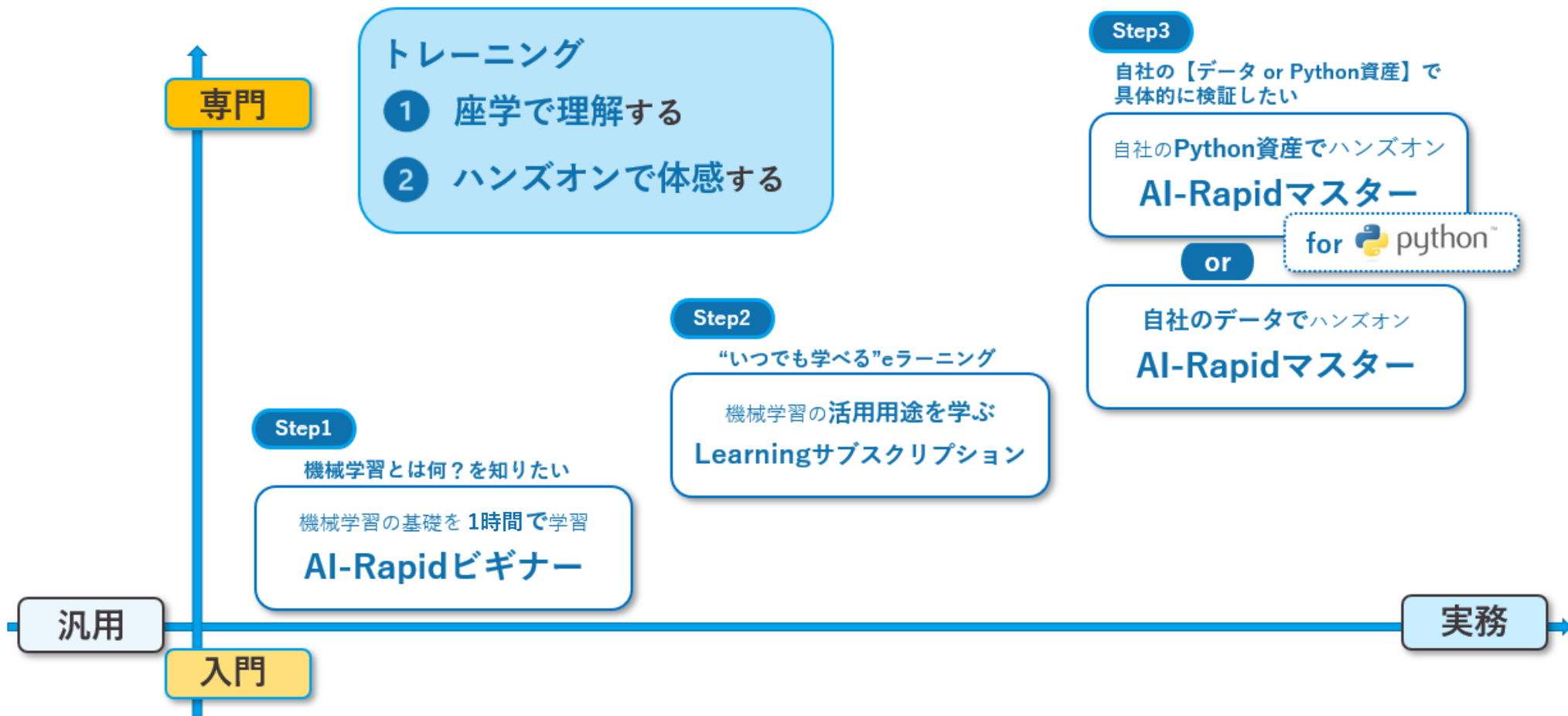
ご質問・お問い合わせはお気軽に  
下記までお願い申し上げます。

[ap-promo@nissho-ele.co.jp](mailto:ap-promo@nissho-ele.co.jp)

日商エレクトロニクス株式会社  
アプリケーション事業本部

# RapidMinerトレーニング情報

AI機械学習ソフトウェア“RapidMiner Studio”フリー版ライセンス



<https://digital-labor.nissho-ele.co.jp/course-rapidminer/>