

ITの、つぎの、幸せへ。



Hirokoの競馬入門 Returns

～ソリューション技術部～

SCSK株式会社

1

▶ はじめに

2

▶ 製品概要

3

▶ AMATERAS RAYによる利益獲得に向けて

4

▶ 最後に

1.はじめに

1.はじめに ～昨年度の振り返り～



上記2製品の特長を生かした観点からどのような分析を行えば良いかを検討。

想定の実分析を実施すべく、それぞれJRA過去データを取り込み、可視化。

完成したダッシュボードを元に、競馬レースの結果を予測し、精度を確認するため、実証。

結果、見事 **的中**！！



1.はじめに

昨年より楽しんで稼ぎたい！！！！



1.はじめに



今年はAIで馬券を
当てるのよ！
by Hiroko

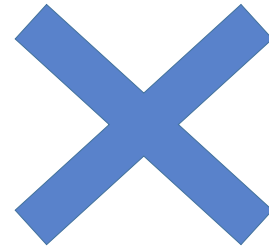
御意



2.製品概要

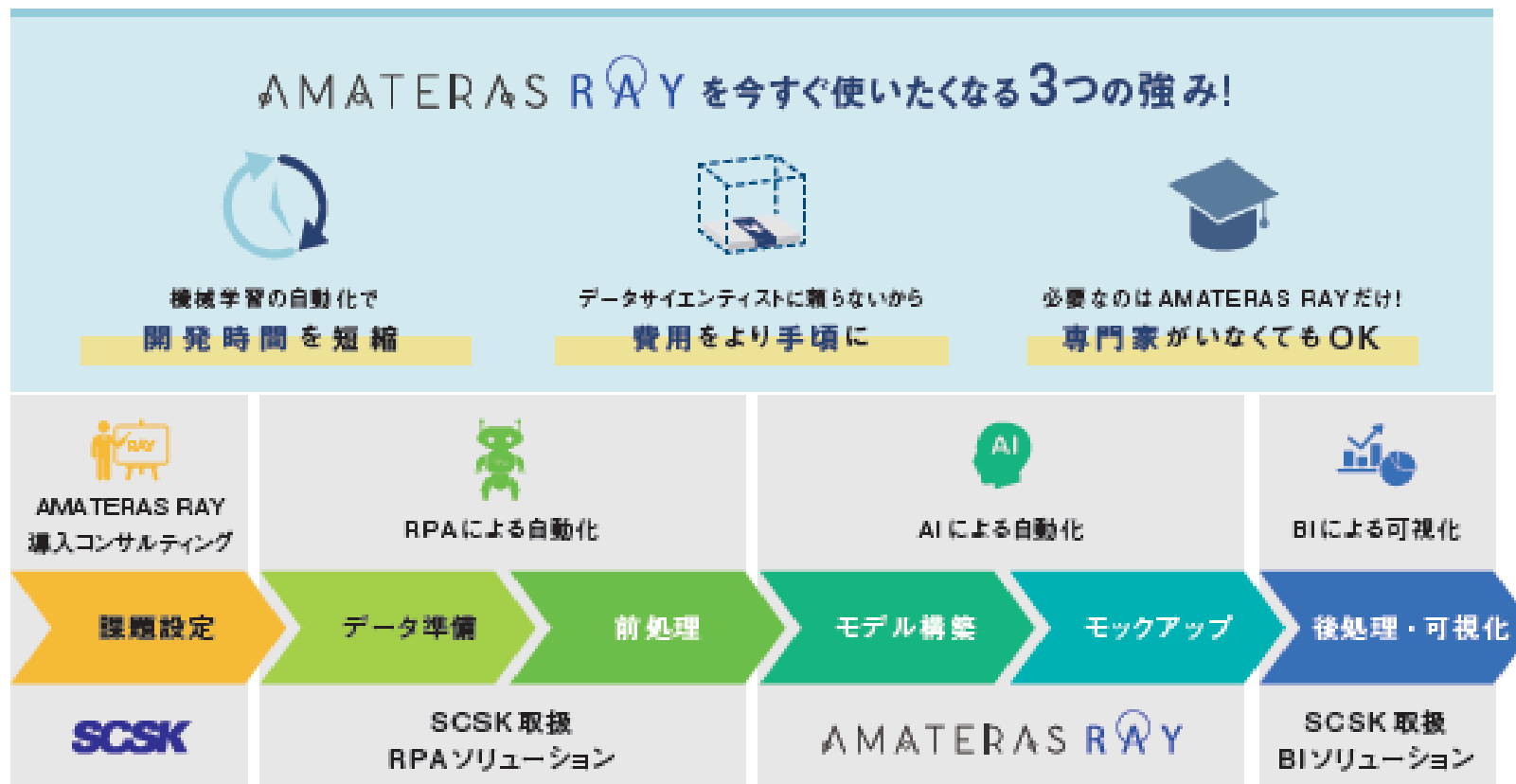
自動機械学習ツール AMATERAS RAY

AIによる競馬予測 × BIによる可視化



2.製品概要

「AMATERAS RAY」は機械学習モデル構築工程を自動化し、
AIの専門家でなくても**簡単にAIによる予測と分類**が可能



2.製品概要

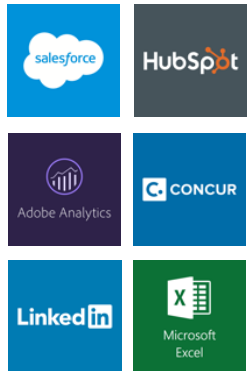
Domoはデータ活用に必要な機能を全て包含する「SaaS」型のサービス

- 必要なデータを1箇所に集約し、ビジネス上の状況をリアルタイムに把握
- コラボレーション機能により効率的なコミュニケーションを通じてアクションを決定
- アラート機能やモバイル対応により継続的なデータ確認を容易に

データを利用したビジネスパフォーマンス向上に必要な機能全てをワンストップで提供

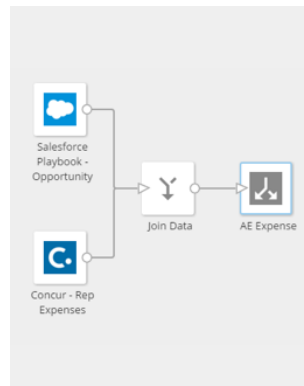
接続

社内外のあらゆるソース
からデータを取り込み



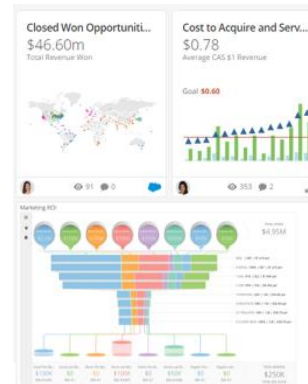
データ準備

クレンジング・結合・整形



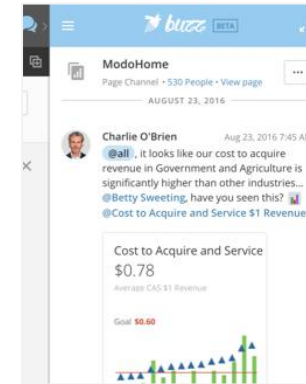
視覚化

データの視覚化
カスタムApp



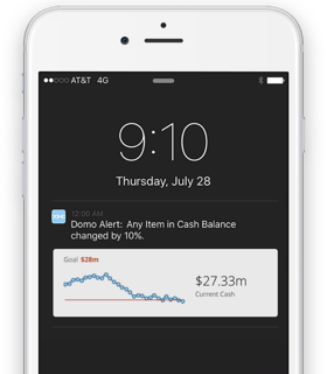
コラボレーション

リアルタイムチャットと
ビジネスコラボレーション



最適化

インテリジェントなアラート
モバイルアプリ



3.AMATERAS RAYによる利益獲得に向けて

3.AMATERAS RAYによる利益獲得に向けて

課題：11月24日 **ジャパンカップ**で馬券を当てたい！

○馬券購入方法

単勝 1着の馬を当てる。

馬連 1着と2着の馬の組み合わせを当てる。

3着まで予測して全組み合わせ(3パターン)を購入

○データはJRAの過去レース結果を使用

データ内容は**DOMO**で可視化。

<https://scsk-jp-partner.domo.com/page/987446684>



3.AMATERAS RAYによる利益獲得に向けて

Home

AMATERAS RAY



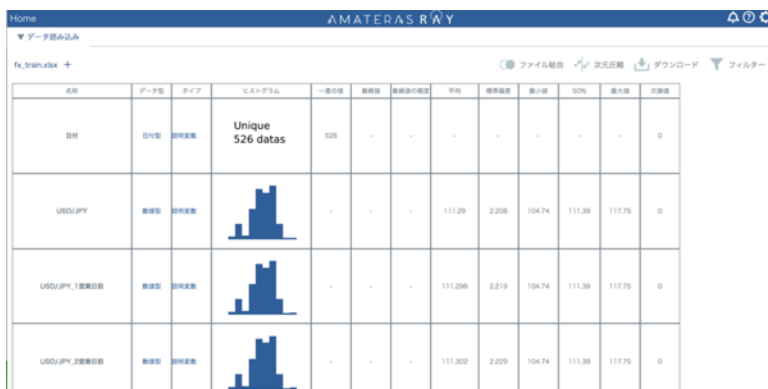
▼ データ読み込み

▼ AIモデル構築

▼ 予測シミュレーション

fx_train.xlsx

たった3つのStepでAIによる予測ができる！



Step1

学習用データの読み込み
と目的変数選択
(予測、分類)



Step2

学習済みモデルの構築
(12アルゴリズム)



Step3

学習済みモデルの選択
と未知データ読み込み
による予測

3.AMATERAS RAYによる利益獲得に向けて

①データ読込

②AIモデル作成

③予測

名称	データ型	タイプ	ヒストグラム	一意の値	最頻値	最頻値の頻度	平均	標準偏差	最小値	50%	最大値	欠損値
騎手名	文字型	説明変数	Unique 798 datas	798	-	-	-	-	-	-	-	1
斤量	数値型	説明変数		-	-	-	54.989	1.807	47	55	66	1
頭数	数値型	説明変数		-	-	-	14.708	2.307	4	16	18	1
馬番	数値型	説明変数		-	-	-	7.854	4.441	1	8	18	1
確定着順	数値型	目的変数 予測 目的変数 分類 使用しない 目的変数 予測		-	-	-	7.727	4.455	-	8	18	1

・ JRAのHPからダウンロードした過去のレース結果をAMATERASにアップ

・ 画面をクリックして説明変数と目的変数を選択

目的変数：[確定着順]

説明変数：[レース名][クラス][トラック][距離][馬場][馬名][父馬名][母馬名][性別][年齢][騎手][斤量][頭数][馬番]

3.AMATERAS RAYによる利益獲得に向けて

①データ読込

②AIモデル作成

③予測

正の相関 上位30

1位	母馬名	0.42
2位	騎手名	0.29
3位	頭数	0.27
4位	父馬名	0.17
5位	レース名	0.15
6位	性別	0.09
7位	馬番	0.09
8位	場所	0.06
9位	年齢	0.02
10位	芝・ダ	0.01

負の相関 上位30

1位	斤量	-0.11
2位	距離	-0.08
3位	クラスコード	-0.02
4位	トラックコード	-0.02
5位	レース番号	-0.01

	確定着順	場所	レース番号	レース名	クラスコード	芝・ダ	トラックコード	距離	馬場状態	馬名	性別	年齢	騎手名	斤量	頭数	馬番	父馬名	母馬名
確定着順	1	0.06	-0.01	0.15	-0.02	0.01	-0.02	-0.08	0	0	0.09	0.02	0.29	-0.11	0.27	0.09	0.17	0.42
場所	0.06	1	-0.01	0.06	0	0.02	-0.02	0.04	0	-0.01	-0.01	0.03	0.07	0.04	0.24	0.07	0.03	0.05
レース番号	-0.01	-0.01	1	-0.13	0.57	-0.21	0.08	0.02	-0.01	0	-0.1	0.56	-0.1	0.19	-0.06	-0.02	-0.11	-0.21
レース名	0.15	0.06	-0.13	1	-0.13	0.18	-0.14	-0.38	0.01	0	0.11	-0.12	0.15	-0.23	0.55	0.15	0.05	0.11
クラスコード	-0.02	0	0.57	-0.13	1	-0.16	0.13	0.16	-0.01	0.01	-0.11	0.45	-0.17	0.2	-0.07	-0.02	-0.1	-0.2
芝・ダ	0.01	0.02	-0.21	0.18	-0.16	1	-0.19	-0.21	0.03	-0.01	-0.07	0.01	0.1	0.02	0.05	0.01	0.1	0.06
トラックコード	-0.02	-0.02	0.08	-0.14	0.13	-0.19	1	0.12	-0.01	0	-0.02	0.04	-0.07	0.05	-0.07	-0.02	-0.06	-0.06
距離	-0.08	0.04	0.02	-0.38	0.16	-0.21	0.12	1	-0.01	0.01	-0.19	0.19	-0.17	0.36	-0.27	-0.07	-0.12	-0.13
馬場状態	0	0	-0.01	0.01	-0.01	0.03	-0.01	-0.01	1	0	0	-0.01	0.01	-0.01	0.01	0	0	0.01
馬名	0	-0.01	0	0	0.01	-0.01	0	0.01	0	1	0.02	0	0	-0.01	0	0	-0.01	-0.01
性別	0.09	-0.01	-0.1	0.11	-0.11	-0.07	-0.02	-0.19	0	0.02	1	-0.12	0.1	-0.53	0.11	0.03	-0.02	0.09
年齢	0.02	0.03	0.56	-0.12	0.45	0.01	0.04	0.19	-0.01	0	-0.12	1	-0.04	0.42	-0.05	-0.01	-0.02	-0.1
騎手名	0.29	0.07	-0.1	0.15	-0.17	0.1	-0.07	-0.17	0.01	0	0.1	-0.04	1	-0.27	0.13	0.03	0.25	0.34
斤量	-0.11	0.04	0.19	-0.23	0.2	0.02	0.05	0.36	-0.01	-0.01	-0.53	0.42	-0.27	1	-0.11	-0.03	-0.02	-0.13
頭数	0.27	0.24	-0.06	0.55	-0.07	0.05	-0.07	-0.27	0.01	0	0.11	-0.05	0.13	-0.11	1	0.27	0.07	0.13
馬番	0.09	0.07	-0.02	0.15	-0.02	0.01	-0.02	-0.07	0	0	0.03	-0.01	0.03	-0.03	0.27	1	0.02	0.03
父馬名	0.17	0.03	-0.11	0.05	-0.1	0.1	-0.06	-0.12	0	-0.01	-0.02	-0.02	0.25	-0.02	0.07	0.02	1	0.31
母馬名	0.42	0.05	-0.21	0.11	-0.2	0.06	-0.06	-0.13	0.01	-0.01	0.09	-0.1	0.34	-0.13	0.13	0.03	0.31	1

【補足】

データの相関がヒートマップで確認出来る。

「母馬名」「騎手名」の相関が特に強い。

AMAZING!



3.AMATERAS RAYによる利益獲得に向けて

①データ読込

②AIモデル作成

③予測

▼データ読み込み ▼AIモデル構築 ▼予測シミュレーション

① JRA_DATA2019_test.csv

目的変数：確定着順 予測

戻る



- ・学習アルゴリズムを選択して学習を実行（全てを同時実行可能）
- ・学習後、各アルゴリズムのスコアが表示

2015年以降のレース結果、
25万レコードを学習

3.AMATERAS RAYによる利益獲得に向けて

①データ読込

②AIモデル作成

③予測

▼データ読み込み ▼ AIモデル構築 ▼ 予測シミュレーション

① JRA_DATA2019_test.csv

戻る

学習モデル

	3.0989 LightGBM 説明
	3.1419 Ridge 説明
	3.1419 BayesianRidge 説明
	3.2772 RandomForest 説明
	3.7831 ElasticNet 説明
	3.7831 Lasso 説明

予測データ

JRA_JC_20191124.txt
2019-11-24T03:15:16

予測結果データ

result_LGBMRegressor_JRA_JC_20191124.csv
2019-11-24T03:16:40

予測する

- ・ 予測データ(JC)をアップロード
- ・ 学習モデルと予測データを選択し、
予測開始
- ・ 予測結果はCSVでダウンロードが可能

精度の高い順にランキング表示してくれるので、
トップにある「LightGBM」を選択。

3.AMATERAS RAYによる利益獲得に向けて

投票結果

購入限度額
購入可能件数

0円
2,694件

続けて投票する

お客様の投票を受け付けました。

受付番号： 0001 受付時刻： 12:59 受付ベット数： 6

印刷する

件数 状況	投票内容	組数	金額	馬券表示
1	東京（日） 11R 単勝	08	1組 500円	
2	東京（日） 11R 馬連	08 - 09	1組 200円	
3	東京（日） 11R 馬連	08 - 15	1組 200円	
4	東京（日） 11R 馬連	09 - 15	1組 200円	
5	東京（日） 11R 単勝	09	1組 200円	
6	東京（日） 11R 単勝	15	1組 200円	
合計金額： 1,500円				

印刷する

ログアウト

ジャパンカップ(11/24)予測結果

AMATERAS 予測順位	馬番:馬名	人気
1	08:レイデオロ	1
2	15:ジナンボー	1 0
3	09:ルックトゥワイス	7

TARGET 予測順位	馬名	人気
1	ワグネリアン	2
2	シュヴァルグラン	8
3	レイデオロ	1

競馬分析ソフト「TARGET」も過去のデータからAIを使用して予測している・・・どちらが当たるのか！？

もちろんAMATERASを信じて馬券**1500円**購入！

3.AMATERAS RAYによる利益獲得に向けて ～結果～

AMATERAS 予測順位	馬名	人気	レース 結果順位
1	レイデオロ	1	<u>1 1</u>
2	ジナンボー	1 0	<u>1 3</u>
3	ルックトゥワイス	7	<u>1 0</u>

TARGET 予測順位	馬名	人気	レース 結果順位
1	ワグネリアン	2	<u>3</u>
2	シュヴァルグラン	8	<u>9</u>
3	レイデオロ	1	<u>1 1</u>

着	枠	馬印	馬名	斤量	騎手
1	3	5	スワーヴリチャード	57.0	マーフィ
2	1	1	カレンブーケドール	53.0	津村明秀
3	2	2	ワグネリアン	57.0	川田将雅
4	8	14	マカヒキ	57.0	武豊
5	4	6	ユーキャンスマイル	57.0	岩田康誠
6	4	7	ダイワキャグニー	57.0	石橋脩
7	7	13	エタリオウ	57.0	横山典弘
8	3	4	ムイトオブリガード	57.0	ルメール
9	6	11	シュヴァルグラン	57.0	スミヨン
10	5	9 ▲	ルックトゥワイス	57.0	デットー
11	5	8 ◎	レイデオロ	57.0	ビュイッ
12	2	3	ウインテンダネス	57.0	田辺裕信
13	8	15 ○	ジナンボー	57.0	ムーア
14	6	10	ダンビュライト	57.0	松若風馬
15	7	12	タイセイトレイル	57.0	M. デム

惨敗！
(1500円→0円)



3.AMATERAS RAYによる利益獲得に向けて ～考察～

負けたままでは終われないので、有馬記念で**倍返し**だ！！

■ 振り返り(仮説)

学習モデルの目的変数に「**確定着順**」を使用していたが、
レースごとに頭数の異なる競馬の予測に着順は向いていないのではないだろうか？

■ 学習モデルのブラッシュアップ(検証)

「着順の早い馬＝速く走る馬」であるため、着順の予測ではなく「**速度**」を目的変数にしてみよう。

■ 馬券購入方法の見直し

どうしても当てたいので単勝、馬連に加えて「**複勝**」を追加。

複勝 3着までに来る馬を一頭予測



3.AMATERAS RAYによる利益獲得に向けて

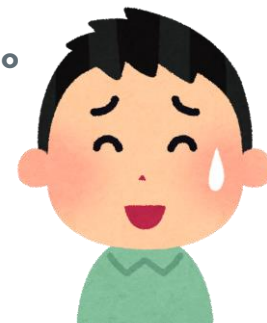
有馬記念(12/22)予測結果

当日分 JRA 受付番号：0001 受付時刻：13:42 受付ベット数：9					印刷する
件数	投票内容		組数	金額	馬券表示
1	中山（日）11R 複勝	06	1組	100円	
2	中山（日）11R 複勝	10	1組	100円	
3	中山（日）11R 複勝	15	1組	100円	
4	中山（日）11R 単勝	06	1組	100円	
5	中山（日）11R 単勝	10	1組	100円	
6	中山（日）11R 単勝	15	1組	100円	
7	中山（日）11R 馬連	06 - 10	1組	100円	
8	中山（日）11R 馬連	06 - 15	1組	100円	
9	中山（日）11R 馬連	10 - 15	1組	100円	
購入金額：900円 払戻金額：0円					

AMATERAS 予測順位	馬番:馬名	人気	オッズ
1	06:リスグラシュー	2	6.7
2	15:アエロリット	12	107.1
3	10:サートウルナーリア	3	7.8

圧倒的1番人気の現役最強馬

オッズ1.5倍のアーモンドアイがいないけど、
AMATERASくん・・・ほんとに？
もちろん信じて馬券900円購入。



3.AMATERAS RAYによる利益獲得に向けて

有馬記念(12/22) レース結果

当日分 JRA 受付番号：0001 受付時刻：13:42 受付ベット数：9					印刷する
件数	投票内容		組数	金額	馬券表示
1 的中	中山 (日) 11R 複勝	06	1組	100円	
2 的中	中山 (日) 11R 複勝	10	1組	100円	
3	中山 (日) 11R 複勝	15	1組	100円	
4 的中	中山 (日) 11R 単勝	06	1組	100円	
5	中山 (日) 11R 単勝	10	1組	100円	
6	中山 (日) 11R 単勝	15	1組	100円	
7 的中	中山 (日) 11R 馬連	06 - 10	1組	100円	
8	中山 (日) 11R 馬連	06 - 15	1組	100円	
9	中山 (日) 11R 馬連	10 - 15	1組	100円	
				購入金額：900円 払戻金額：4,140円	

オッズ30倍！

AMATERAS 予測順位	馬名	人気	結果
1	06:リスグラシュー	2	1
2	15:アエロリット	12	14
3	10:サートウルナーリア	3	2

的中！
(900円→4140円)

ジャパンカップ分も取り戻し、トータルで勝利！

3.AMATERAS RAYによる利益獲得に向けて



今年も利益を出した
なんてさすがだわ！
さすが私の部下！
褒めてつかわす！
by Hiroko

有りがたきお言葉



4.最後に

これまでの“人”による予測に比べ、“AI”により簡単に正確な予測が可能ですが、それは質のよいAI（機械学習）モデルが作成できてこそとなります。

AI（機械学習）のモデルは一度作ったら終わりではなく、検証と調整の繰り返しが大切であり、その為には、簡単に誰でもモデルを作成可能な使いやすいツールが必要となります。

