RapidMiner Studioでモデル運用

-Deployment機能紹介-

株式会社KSKアナリティクス





モデル作成の先へ

RapidMiner Studioで様々なモデルを作成し、 効果的な予測や洞察を得ても、それは一時的なものになりがちです。 しかし、実際の業務の大半は繰り返し続いていくものです。 よって、業務と同じようにモデルを継続モニタリングすると、 ビジネスインパクトをより大きくすることが出来ます。

本資料では機械学習モデルを本番運用していくために、

AutoModelとDeploymentの連携機能を紹介していきます。

<u>※DeploymentはRapidMiner Studio Enterprise版の機能です</u>



Deployment概要

Deploymentには以下のような機能があります。

- ・モデルを整理し、重要なデータを一つの場所に保管する
- ・モデルのパフォーマンスを経時的に追跡し、ドリフトやバイアスの警告が可能
- ・共通のプロジェクトで協力しているグループで共有することが可能
- ・webサービスを通じて、他のアプリケーションと統合することが可能

以下の表はどのような場合にどの機能が使えるのかを示しています。 今回はRapidMiner Studio + DB接続で使える機能をご紹介します。 ※DB接続は環境によって異なりますので、詳細はDB管理者にご確認ください。

	RapidMiner Studio	+データベース接続
RapidMiner Studio	•Models •Simulator •Scoring	•Dashboard •Performance •Drifts
+ RapidMiner Al Hub	 Integrations 	•Alerts



Locationの作成

DeploymentにはAutoModelからすぐ飛べますが、 初めての方はLocationを設定しておく必要があります。 デザイン画面でDeploymentタブを押し、案内に沿ってNEWボタンを押します



Locationの作成

Locationの名前を設定したら、LocalかRemoteか選択します。 Remoteの利用にはRapidMiner AIHubが必要です。 ここではLocalを選択します。

保存するリポジトリを聞かれたら、新規で空フォルダを作り、選択します





Locationの作成

dhan a r

このまま、作成完了でも良いですが、Monitoringを使いたいので、 Activate Monitoring?(赤枠)をチェックします。 DB接続が無ければ、新規作成します。

w New Deployment Location						
	Name	Location	Storage Folder	Monitoring	Alerts	
You may disable monitoring of deple champion-challenger-tests. You nee This connection must be to a databi as well as for querying them. Only M database (for remote locations) do r	oyed models. If ena ad to specify a conr ase in which all log ySQL and Postgres not need to be the s	abled, however, y nection object wh data will be stor s are supported same.	rou will be able to track lich is located in the sa ed - hence you need to as database for model	model performance, me repository than th have the necessary a monitoring. Please n	input drifts, and perform e folder you have select access rights for creating tote that the monitoring o	ad on the previous scree a tables and data entries latabase and the Al Hub
Select Database Connection:	sample //sample/Connecti	ions/sample				
	Refresh Conner	ctions Add	New Connection	Show Connection	Info	
			And the second sec			



Deploymentの作成

Locationが作成できたら下図のADD DEPLOYMENTをクリックします 次に、Nameと対象の分析が分類か回帰なのかを入力します。





Deploymentの作成

Deploymentが作成できました。 このDeploymentにモデルを追加していきます。 AutoModelから簡単に追加できますので、 AutoModelでお好きなモデルを作成してください。

Overview Shows all deployments for the selecte	d deployment lo	cation. (j)				NEW CONN	ECT MANAGE	test //sample/demo	•
Total Scores O	Activ	0 Deployments	Week	ly Scores O	Aler O	ts			
Deployment	-	Explain	Number of Models	Model Type	Alerts	Last Month Activit	у		
test		•	0	Classification	0				



AutoModelとの連携

下図はAutoModelの結果からRundom Forestのモデルを表示しています DEPLOYをクリックすると、配置先を聞かれますので、 Nameを入れ、先ほどのLocation、Deploymentを選択します





AutoModelとの連携

同じようにモデルを複数配置してみました。 この画面では右クリックで各モデルの概要を確認できるほか、 チャンピオンとチャレンジャーを設定できます。 StatusがActiveのものがチャンピオンモデルです。

Lest: Models Shows all models in this deployment and allows to activate models or use them as challengers. (i) DASHBOARD MODELS PERFORMANCE DRIFTS SIMULATOR SCORING ALERTS INTEGRATIONS DEPLOY AUTO MODEL DEPLOY CUSTOM MODEL DEPLOY CUSTOM MODEL Deploy Custom Model									
Name	Туре	Created	Author	Status	Predeployment Error				
Random Forest	Random Forest	2021/02/03 16:05:59	Unknown	Active	21.1% ± 5.2%				
Decision Tree	Decision Tree	2021/02/04 14:18:15	Unknown	Challenger	15.6% ± 7.6%				
Logistic Regression	Logistic Regression	2021/02/04 14:17:52	Unknown	Challenger	50.0% ± 3.2%				

Dashboard

チャンピオンモデルの結果の概要や推移を確認できます。 グラフの表示幅は3か月、1か月、1週間、1日と変更できます。





Performance

Dashboard機能よりもさらに詳細に結果情報を確認できます。 チャンピオンとチャレンジャーの比較も見ることが可能です。





Drifts

ドリフト(漂流)は学習時からの時間経過や状況の変化によりデータの特徴が 変化したことで予測誤差が増加する現象を指します。 ドリフトが大きくなりすぎたらモデルを再度学習させる必要があるので、 こちらの画面でドリフトの状態を確認しましょう。

On-	time Arrival Forecasting: Drifts	Active? Back.to.Overview
DASHE	OARD MODELS PERFORMANCE CRIPTS SIMULATOR SCORING ALERTS INTEGRATIONS	
Drift fo	r Factors	Distributions for LateAircraftDelay
	150X	0.014
	50X	0.012
	sox	
	70%	60
	40%	
	5 JON	
	40%	£ 0005
	30%	5004
	20% Landworad/Delay	
	1C Drift ES	6.652
	71/11//////////////////////////////////	0 -75 -50 -25 0 25 50 75 100 125 150 Tealming, Scaring
	✓ Factor	Training Scoring



Simulator

AutoModelで提供されているSimulatorと同機能です。 詳しい説明はここでは割愛しますが、Optimizeも付いていますので、 AutoModelの時と同じように利用できます。



Change inputs to the model and see its predictions in real time.

DASHBOARD MODELS PERFORMANCE DRIFTS SIMULATOR SCORING Most Likely: No Age: 30.088 1 100% 90% 0.385 ٢ No of Parents or Children on Board: 80% 70% 60% 50% No of Siblings or Spouses on Board: 0.525 ٢ 40% 30% 20% 10% Passenger Class: Third • 1 0% Yes No 32.476 1 Passenger Fare: Important Factors for No ٢ Port of Embarkation: Southampton V Sex Sex: Male v V Passenger Fare What is this? Optimize Passenger Class



Back to Ov

Scoring

De

モデルに新しいデータを適用し、結果を確認できます。 Monitoringを有効にしている場合、この結果が記録され、 DashboardやPerformanceで推移を確認することが出来るようになります。 また、結果はリポジトリやExcel、CSVに出力できます。

edit	Risk T	rial: So	coring	× O													Back to
HEOARD	MODELS	PERFORMANCE	DRIFTS	SIMULATOR	SCORING	ALERTS 10											
ORE DAT	A DEFINE A	CTUALS															
ror Rat	e: 23.79%																
w No.	creditworthy	prediction(c	coefidence(_	confidence(_	cost	status	credithistory	purpose	savings	employment_	property	otherinstal.	housing	job	phone	gender	duration
	Worthy	Worthy	0.172	0.828	0.657	+ 0 Euro	offical account	radoblevision	unincentine	>=7 years	real estate	none	own	skilled emplo	yes, registere	male	6
	Not Worthy	Worthy	0.395	0.605	0.209	< 200 Euro	existing credit	radiofelevision	+ 100 Euro	1-12_<438.	real estate	none	own	skilled emplo	none	female	48
	Worthy	Worthy	0.170	0.830	0.661	no checking	offical account	education	_ + 100 Euro	4+=_+738_	real estate	none	own	unskilled-re	none	male	12
	Worthy	Worthy	0.333	0.667	0.333	<0 Euro	existing credit.	fumiluraireau	_ < 100 Euro	4<=_<718_	building soci	none	forfree	skilled emplo_	none	male	42
	Not Worthy	Not Worthy	0.522	0.478	0.045	<0.Euro	delay in payin.	car (new)		1-1-418	unknown/ns	none	for thee	skilled emplo	none	male	24
	Worthy	Worthy	0.262	0.738	0.476	no checking .	existing credit	education	unknown/no	1-=	unknown/no	none	forftee	unskilled-re	yes, registere	male	36
	Worthy	Worthy	0.183	0.817	0.634	no checking	existing credit.	fumburelegui	500 <= _ < 1_	_>=7years	building soci	none	own	skilled emplo.	none	male	24
	Worthy	Worthy	0.304	0.695	0.392	< 200 Euro	existing credit.	car (used)		1-0.418.	car or other	none	rent	management.	yes, registere	male	36
	Worthy	Worthy	0.165	0.835	0.669	no checking	existing credit.	radioNelevision		4+x_+7j8_	real estate	none	040	unskilled-re_	1014	male	12
	Not Worthy	Worthy	0.424	0.576	0.153	< 200 Euro	critical account	car (new)		unemployed	car or other	none	own	management.	none	male	30
	Not Worthy	Not Worthy	0.605	0.395	0,210	<0 Euro	existing credit.	business	< 100 Euro	_<1year	building soci	0004	rent	skilled emplo	none	female	48
	Worthy	Worthy	0.195	0.805	0.509	< 200 Euro	existing credit	radiofelevision		1-1-4 18	car or other	none	own	skilled emplo	yes, registere	female	12
	Not Worthy	Worthy	0.340	0.660	0.319	< 0 Euro	offical account	Car (Mex)	< 100 Esro	.×=7years	car or other	none	own	unskilled-re	0054	male	24
	Not Worthy	Worthy	0.338	0.662	0.325	<0 Euro	existing credit	radiofelevision	100 += _ +5	1-0	car or other	none	own	unskilled-re	none	female	24
	Worthy	Not Worthy	0.549	0.451	0.098	<0 Euro	No credits tak	business	and the second second	_<1year	car or other	bank	own	skilled emplo	none	male	30
	Not Worthy	Not Worthy	0.554	0.445	0.109	< 200 Euro	existing credit.	car (used)			unknown/no_	none	foritee	management.	yes, registere	female	24
	Worthy	Worthy	0.185	0.815	0.629	no checking	existing credit.	radiofelevision	500 -= 1	≫7years	car or other	none	own	skilled emplo	yes, registere	male	24
	Worthy	Worthy	0.199	0.801	0.602	no checking	offical account	Car (New)	. + 100 Euro	1-0	car or other	none	own	skilled emplo	yes, registere	male	9
	101-march	Window.	A #04	A.85.4	A 255			and a Reference of		100 120	sand and de			shifts d smale		mate	



AI Hubとの連携

AI Hubを導入すると、以下の機能も追加で利用が可能です。

• Alerts

DriftsやPerformanceの結果に対して、トリガーとなる数値を設定し、 その数値からはみ出した際にAlertを出せるようになります。

・Integrations 結果をWebサービスに統合して、別のアプリケーションで使用できます。

On-time Arrival Forecasting: Alerts							REMOTE Monitored Deployment: Integrations Merri C Integrations Int				
Alerts CREATE NEW ALERT		14					You can use the URL below for integrating the also test the senice below to see a full exampl Scoring URE: http://scalhost.8885/u	predictors of the achier model and simple applications. You will need to authenticate on the server. The query for a size of a the resource is thin the same. parentstprocesss/remote_monitored_deployment_score	nd is the URL show blow blow blow by the any number of party parameters defining the values of the equilibrium, i.e. the formula VRL/totari valuet Middad valued K. * The can () (by discalant 2000 generatives values are made or parameter, blow, unlaster)		
10	Trigger	Value	Timeframe	Check Frequency	Quiner	Email					
Drift Checker	diff praterban	0.500	Last Reek	Every Day	DEFAULT		Test weatherdelay	4.300151400454201	Teat URL A http://ocateat.illull/appliest/process/inmole_monthmed_deployment_score/weatherdelay=4.30015140045420156atearcatibelay=25.750037850113556crs.ardime=19		
Error Rate Checker	ang emir prester than	1.03	LastDer	EveryHour	DEFHILT		lateansrafidelay	25 75003785011385	885 X 12 Cm-C		
Scoring Volume	less screet than	1000	Latting	Every Day	DEFAULT		orsantime	1622.7882	The Dev		
Time to chaine models?	achie model worse than challenger	1	Latiber	EastHow	DEFAULT	taning@spotenter.com	distance	733 9729	X Delete Deem		
							securitydelay	0 13217259651778856	C market Arrive		
							origin	040			
							orsdeptime	1455.65375			
							nasdelay	13.498107494322483			
Notifications							dest	ORD	TestRespone		
Nouncations							dep-detay	43.5851	Response time and ma		
ACKNOWLEDGE ALL							uniquecamer	90%i	ComparCanter "VM", "DarOMeet" 3 86745, "DepTime" 1511 65085, "ORDepTime" 1455 65375, "ORDerTime" 1522 7892 "Flightham" 2332 255, "DepDeta" 43, 6651," Distance" 233 8728, "TaxOut" 17 2005, "CanterCenter" 19 08115064345193, "MADuale" 13 48157248422483, "LatercentDetary" 25 750378051555, "Solider" and "Adu Contence" 233 8728, "TaxOut" 17 2005, "CanterCenter" 19 08115064345193, "MADuale" 13 48157248422483, "LatercentDetary" 25 750378051555, "Solider" and "Adu		
0	Count	Lest		neson			fightrum	2332,252	alCen" mult "confidence/bir/10.753886075847264" confidence/tex/10.246103824152736" spedictoryChTeneAetval/" 76v" (cost" 0.5077821516645281)		
Environment 1 Fill Star 2012/2010 Aug. Environia 252% in the last 1 days.							deptime	1511.65085			
Scoring Volume	3	Fit Sep 27	151140 CERT 2019	lumber of scores has been 759 in the last 1 days.			taniout	17.72825			
							daysfereri.	3.96745			
							TEST WEB SERVICE				



モデル運用しよう!

今回はRapidMiner Studioだけで出来るモデル運用のための機能紹介でした。 RapidMinerを使えばモデルの作成だけでなく、

パフォーマンスの推移やドリフトの発生も分かりやすく確認できるので、 運用が非常に楽であることがお判りいただけたと思います。

更にAI Hubとの連携も行えば、Deploymentをチーム間での共有や、 他のアプリケーションとの連携、Alert機能も使えるようになりますので、 チーム体制での運用が効率的に実施できます。 是非、RapidMinerでモデル運用をお楽しみください!

