

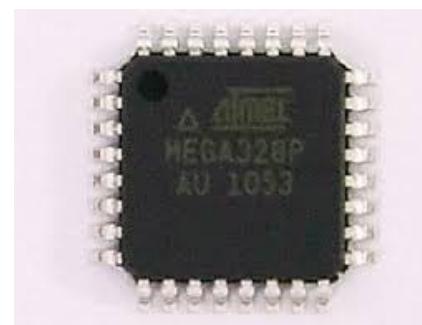
JAWS DAYS 2015

IoTデバイスとクラウドを使って センサーハック～機械学習～

株式会社プリファードインフラストラクチャー

自己紹介

- 名前：河原 一哉（36：年男）
- 所属：製品事業部
- 好きなOS：Solaris
- 好きなマイコン：ATMEGA3238P-AU
- 好きな護衛艦：ごくりゅう





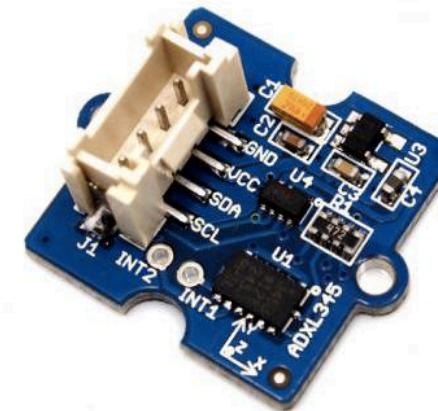
よろしくおねがいします！

ハンズオンのシナリオ

- 今日の目標
 - 加速度センサーの値を使ってPCのキーボード入力を監視しよう
- 今日の作業
 - シナリオ 0 : センサーからデータを取得してみる
 - シナリオ 1 : 学習データを作成して精度測定する その 1
 - シナリオ 2 : 学習データを作成して精度測定する その 2
 - シナリオ 3 : 実際に分析を実施してみる
 - センサーデータをkinesisに送る
 - センサーデータをkinesisから受け取る

ハンズオンのシナリオ

- 今日使うセンサー
 - 3軸加速度センサー ADXL345
 - I2C通信対応
 - 電源電圧：3.0V ~ 5V DC
 - 待機時消費電力 : 0.1 μ A



ハンズオンのシナリオ

- 今日使う環境
 - Intel : Edison
 - Amazon : EC2, Kinesis
 - PFI : Sedue for BiFata, Sedue Predictor



Amazon EC2



準備

- 今日使うプログラムとかのダウンロード
 - ここからダウンロード：
 - https://github.com/kawaii/jaws2015_hands_on
- AWS環境の準備
 - **JAWS DAYS 2015 HackDayAWS側構築ハンズオンテキスト**
 - HackDay-Track3-Script-Rev0.1.docx
 - 動作確認：<http://hostname>

今日使うPFI製品

- Sedue for BigData
 - リアルタイム大規模データ分析基盤
 - 高速で漏れの無い全文検索が得意
 - 今日はリアルタイムな部分だけを使います。
- Sedue Predictor
 - オンライン機械学習プラットフォーム
 - 特長抽出、学習、分類、精度評価を簡単に
 - 今日はこっちを主に使います。



機械学習とは？

- 機械学習とは
 - 「明示的にプログラミングすることなく、コンピュータに行動させようとする科学」
 - 学習により、自動的にルールを見つけ出し、分類や予測などを行なえるようになる機能
- 学習の種類
 - 教師あり学習：人間が明示的に正解を教える
 - 教師なし学習：明示的な正解は与えない
 - 半教師あり学習、教科学習とかも。。。。

機械学習とは？

- 機械学習を使うために必要な事
 - データを作成する
 - 学習データを作成する
 - 特徴抽出の方法を考える
 - 学習をする
 - 精度を測定する
 - 機械学習を使用する

機械学習を使ってなにをするか？

- 機械学習はいろんな事ができる
 - 分類：カテゴリに分ける
 - 回帰：数値を予測する
 - クラスタリング：いくつかに分ける
 - 異常値検知：異常値を見つける
- 今回の問題設定
 - 加速度センサーの値を読み取って、以下の状態に分類してみる
 - キー入力している状態
 - キー入力していない状態

シナリオ0：センサーからデータを取得してみる

- センサーを調べる
 - ADXL345というチップを搭載
 - ライブラリを探す
 - <https://software.intel.com/en-us/iot/sensors>
- センサーを繋げる
 - 今回のセンサーはI2C接続なので、Dポートに繋げる
- センサーからデータを取得するプログラムを書く
 - サンプルプログラムを見れば大体わかる。
 - ライブラリのおかげで超簡単

注：本日使用するプログラムはここからダウンロード：

シナリオ0：センサーからデータを取得してみる

- Edisonへログイン
 - ssh root@Edisonのホスト名
- awk-sdkのインストール（入ってる人は不要）
 - # node install aws-sdk
- プログラム作成
 - 以下のURLのサンプルを参考に作ってみよう
 - <https://software.intel.com/en-us/node/541816>
- プログラムの実行
 - # node プログラム名
 - 値が取れた！はず？

シナリオ0：センサーからデータを取得してみる まとめ

- シナリオ0の結論
 - センサーデータから値を取るのは簡単
 - ただしライブラリがあれば。無いと結構大変です。
→ ライブラリのあるセンサーを使用しましょうw
(ライセンスには注意してください。)
 - 取れたセンサーの値を見ても何も判らない
 - 可視化しないとなんのこっちゃ判らない。
 - 可視化しても何も判らない
 - 値が上下してるぐらいしか判りません。。。
- 熟練すると判るらしいけど普通は無理。

→ 機械学習を使ってデータの分析をやってみよう！

シナリオ1：学習データ作成、精度測定 その1 Predictorの準備

- Sedue Predictorの準備
 - EC2インスタンスの80番ポートにHTTPアクセス
 - Sedue Predictorをクリック
 - ログイン
 - 使用アカウント : demo@demo.local / jaws2015
 - モデルの作成
 - new modelボタンをクリックしモデルを作成
 - 一緒に1個コンフィグが作られる

シナリオ1：学習データ作成、精度測定 その1 Predictorの選択

The screenshot shows a web browser window with the URL `localhost:8080` in the address bar. The page title is "Sedue製品デモ仮想環境". The content includes a brief description, a "Learn More" button, and three product cards: "Sedue Predictor" (highlighted with a red box), "Sedue for BigData", and "Sedue Manager". A blue callout bubble points to the "Predictor" card with the text "Predictorをクリック".

Sedue製品デモ仮想環境

Sedue製品ファミリを仮想環境上で簡単に体験いただけます。

データの登録はFluentd経由で簡単に行なえます。

Learn More

仮想環境上で体験できる製品

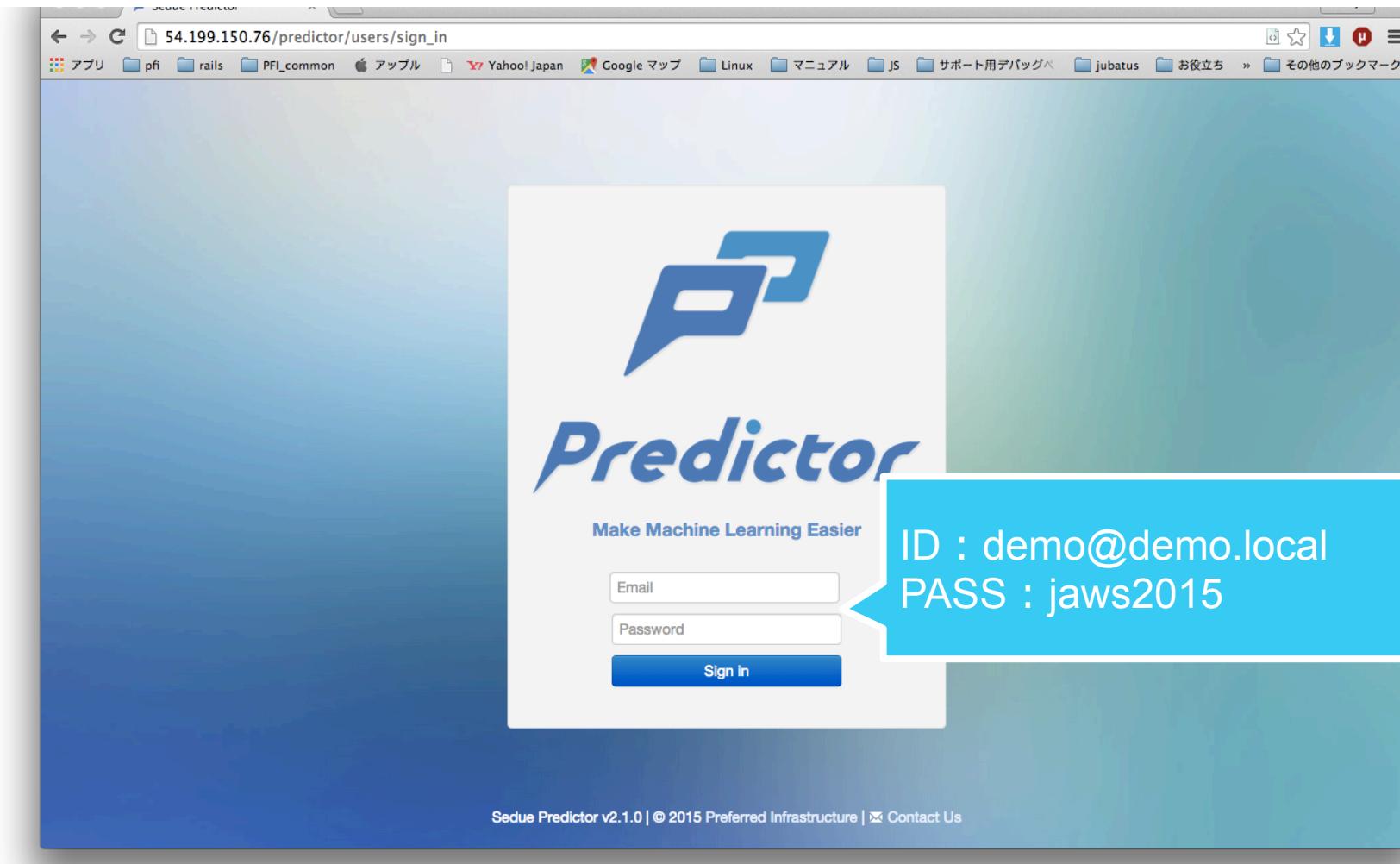
Predictorをクリック


Sedue Predictor
オンライン機械学習システム

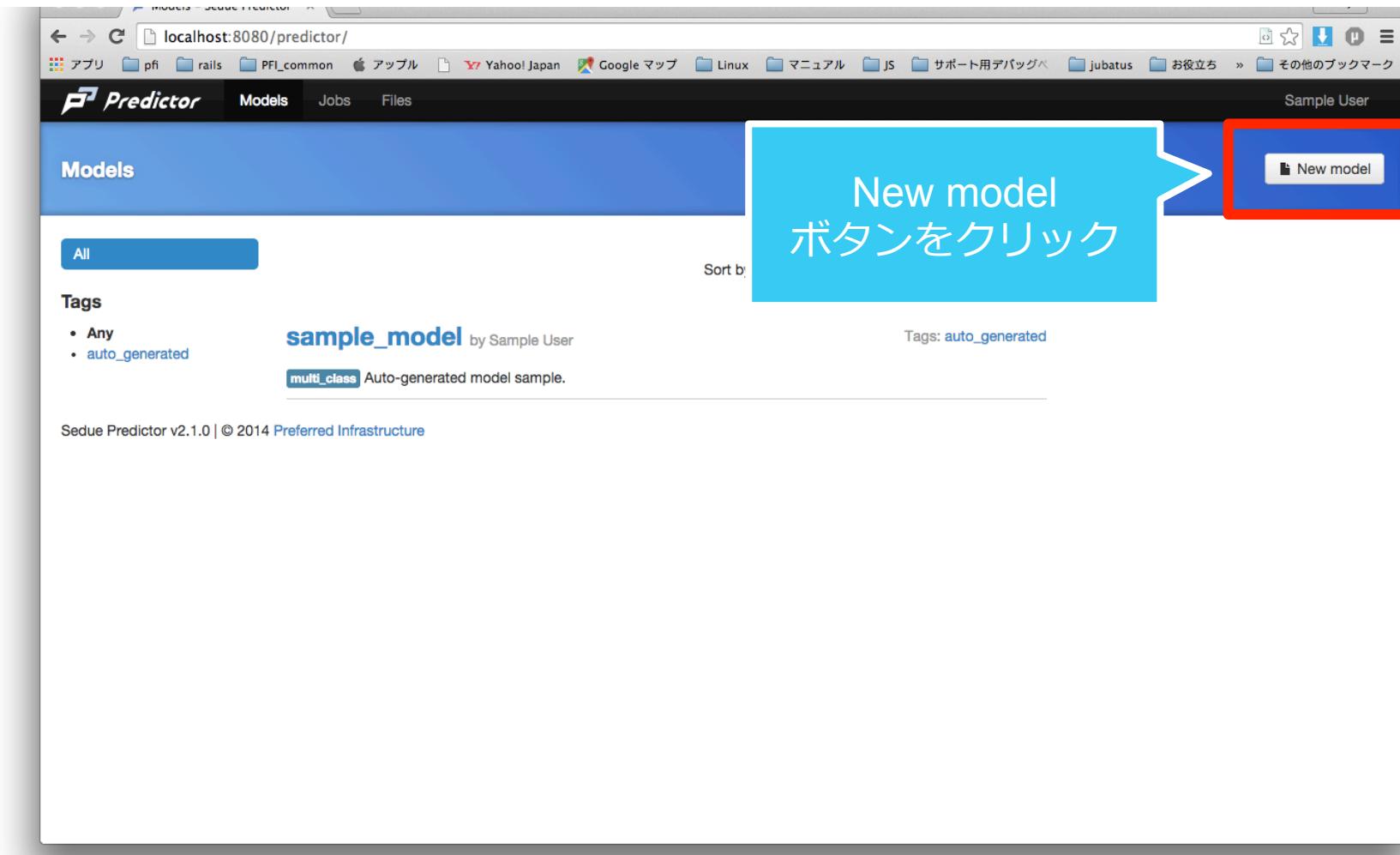

Sedue for BigData
リアルタイム大規模データ分析基盤


Sedue Manager
管理画面

シナリオ1：学習データ作成、精度測定 その1 ログイン画面



シナリオ1：学習データ作成、精度測定 その1 モデルの作成



シナリオ1：学習データ作成、精度測定 その1 モデルの作成

New model

Model

Name

Description

Tags

Model type

Model's config

Name

Description

Easy config Method

Regularization weight

シナリオ1：学習データ作成、精度測定 その1 モデルの作成

New model - Sedue Predictor

localhost:8080/predictor/models/new

Name: default

Description:

Easy config

Method: arow

Regularization weight: 1.0

String rule: space

Num rule: num

Fields: Use all fields
 Select fields

Batch train config

Train method: using_all_data

Train times: 1

Create model

Sedue Predictor v2.1.0 | © 2014 Preferred Infrastructure

シナリオ1：学習データ作成、精度測定 その1 モデルの作成

The screenshot shows the Predictor web application interface at localhost:8080/predictor/models/3/configs. The top navigation bar includes links for Predictor, Models, Jobs, and Files, along with a Sample User dropdown.

Configs tab selected:

ID	Config name	Created at	Default config	Retrain	Last retrain status	Duplicate	Description
3	arow_num_all	15-03-21 18:59:50	★ Default		Not Trained		
4	nherd_num_all	15-03-21 19:00:21	Set default		Not Trained		

Model info section:

Model name (#id)	demo (#3)
Description	(No description)
Model type	multi_class
Owner	Sample User
Tags	

Download training data section:

<input type="checkbox"/> With IDs	Status	Result	With ID	With Predictions	Encoding	File created at						
	mint_documents-v...tar.gz		2015-03-17_1905.zip		2015-03-16_1633.zip		noname		put		すべてを表示	x

シナリオ1：学習データ作成、精度測定 その1

学習データの作成（Edison上の作業）

- 学習データの作成

- センサーの値を1回取得し、その際に、キー入力があったか無かったか正解データを付ける
- PredictorはCSV形式でGUIから学習データを登録できる
- # cd hands_on/edge/session_1
- # node trainer.js > train.data
 - このプログラムはCSV形式でデータを作成します。
 - コンソール上でキー入力（強く叩きすぎないで。弱すぎもダメ）
 - 1000個学習データが出来ると終了。
- ヘッター情報を付ける
- # cat header.txt train.data > train.csv
- ローカルPCにファイルをコピー

シナリオ1：学習データ作成、精度測定 その1 Predictorの設定

- 学習データの登録
 - training dataを押す
 - upload fileボタンをクリック
 - ファイルを選択、アップロード
- コンフィグの作成
 - いろんなコンフィグを作ってみる
- 精度評価
 - Evaluationを選択
 - target configでall configを選択、他はデフォルトでOK
 - submitボタンを押す
 - Accuracyを確認

シナリオ1：学習データ作成、精度測定 その1 学習データの登録

The screenshot shows a web browser window with the URL `localhost:8080/predictor/models/3/configs`. The page title is "demo (#3)". The navigation bar includes links for "Configs", "Training Data" (which is highlighted with a red box), "Annotations", "Fill Annotation", "Evaluation", and "Inspection". A "Sample User" link is also present.

A large blue callout box with the text "Training Data をクリック" (Click Training Data) points to the "Training Data" tab in the navigation bar.

The main content area displays a table titled "Config" with two rows:

ID	Co	Retrain	Last retrain status	Duplicate	Description
3	arc		Not Trained		
4	nhi		Not Trained		

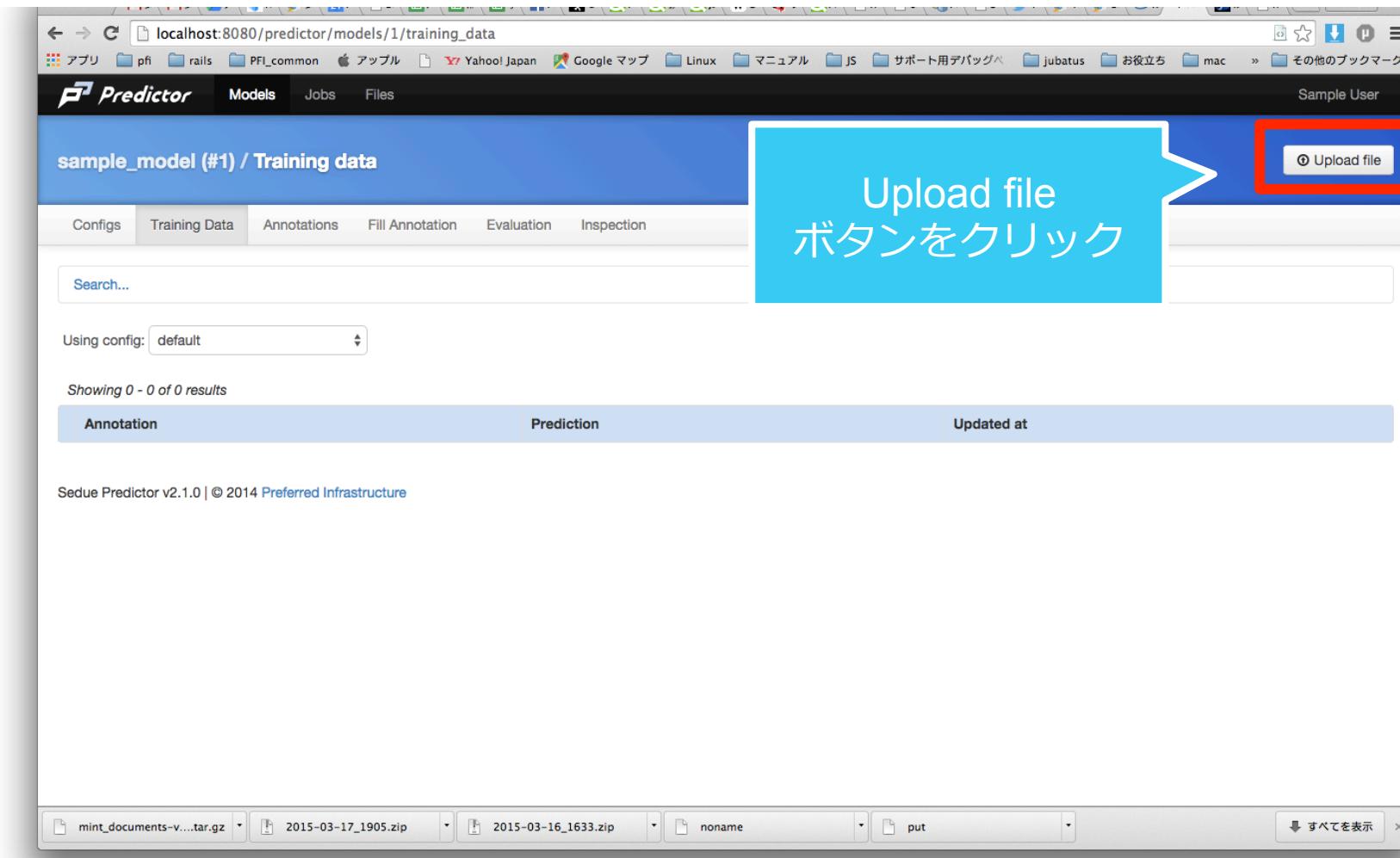
Below this, there is a section titled "Model" containing the following fields:

Model name (#id)	demo (#3)
Description	(No description)
Model type	multi_class
Owner	Sample User
Tags	

At the bottom, there is a section titled "Download training data" with a file list:

Status	Result	With ID	With Predictions	Encoding	File created at
	mint_documents-v....tar.gz		2015-03-17_1905.zip		2015-03-16_1633.zip
			noname		put

シナリオ1：学習データ作成、精度測定 その1 学習データの登録



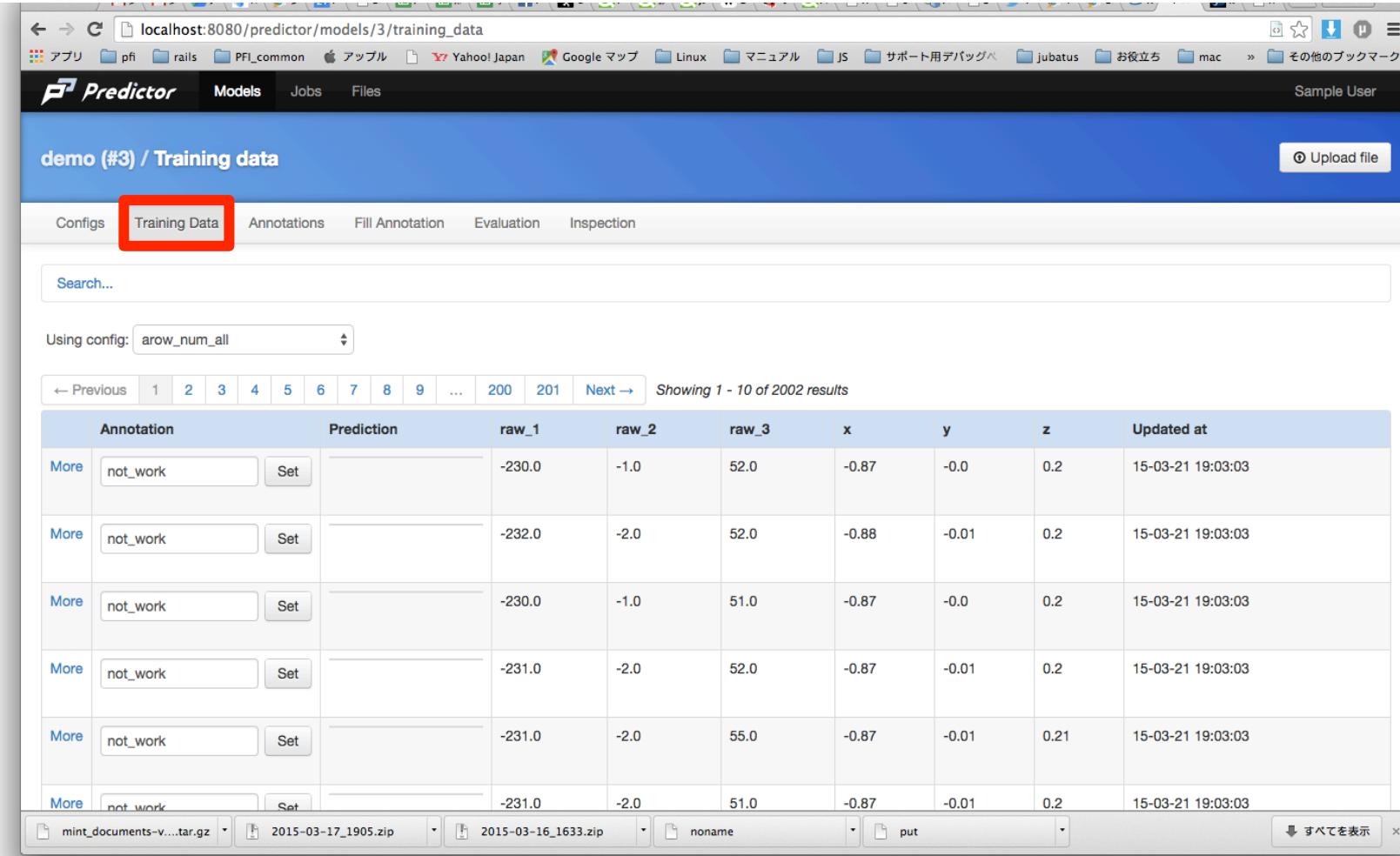
シナリオ1：学習データ作成、精度測定 その1 学習データの登録

The screenshot shows a web browser window for the Predictor application at localhost:8080/predictor/models/1/training_data/upload?config_id=1. The interface is titled "sample_model (#1) / Training data / Upload". A red box highlights the "CSV file" input field, which contains the placeholder text "ファイルを選択 [選択]" (Select file). A large blue callout box with white text says "ファイルを選択して Uploadボタンをクリック" (Select file and click the Upload button). Below the input field is a "Upload" button. At the bottom left, there's a green "SUCCESS" button and the message "200 data added and 0 data updated.". To the right is a table showing upload history:

File	Created at	Started at	Finished at
hands_on5.csv	15-03-19 10:11:21	15-03-19 10:11:24	15-03-19 10:11:27
hands_on5.csv	15-03-16 11:07:44	15-03-16 11:07:46	15-03-16 11:07:50
hands_on2.csv	15-03-16 11:06:55	15-03-16 11:06:56	15-03-16 11:06:57

At the bottom of the browser window, there's a file bar showing several files: "mint_documents-v...tar.gz", "2015-03-17_1905.zip", "2015-03-16_1633.zip", "noname", "put", and a "すべてを表示" (Show all) button.

シナリオ1：学習データ作成、精度測定 その1 学習データの登録



The screenshot shows a web browser window with the URL `localhost:8080/predictor/models/3/training_data`. The page title is "demo (#3) / Training data". The navigation bar includes links for "Configs", "Training Data" (which is highlighted with a red box), "Annotations", "Fill Annotation", "Evaluation", and "Inspection". A "Upload file" button is located in the top right corner. Below the navigation, there is a search bar labeled "Search...". A dropdown menu shows "Using config: arow_num_all". The main content area displays a table of training data results, with the first 10 of 2002 entries shown. The table columns are: Annotation, Prediction, raw_1, raw_2, raw_3, x, y, z, and Updated at. Each row contains a "More" link, an input field for "Annotation" (set to "not_work"), a "Set" button, and various numerical values for the other columns. The "Updated at" column shows the timestamp "15-03-21 19:03:03" for all rows. At the bottom of the table, there are navigation links for "← Previous", page numbers (1 through 201), and "Next →". The status bar at the bottom of the browser window shows several file icons and names: "mint_documents-v....tar.gz", "2015-03-17_1905.zip", "2015-03-16_1633.zip", "noname", "put", and "すべてを表示".

	Annotation	Prediction	raw_1	raw_2	raw_3	x	y	z	Updated at	
More	not_work	Set		-230.0	-1.0	52.0	-0.87	-0.0	0.2	15-03-21 19:03:03
More	not_work	Set		-232.0	-2.0	52.0	-0.88	-0.01	0.2	15-03-21 19:03:03
More	not_work	Set		-230.0	-1.0	51.0	-0.87	-0.0	0.2	15-03-21 19:03:03
More	not_work	Set		-231.0	-2.0	52.0	-0.87	-0.01	0.2	15-03-21 19:03:03
More	not_work	Set		-231.0	-2.0	55.0	-0.87	-0.01	0.21	15-03-21 19:03:03
More	not_work	Set		-231.0	-2.0	51.0	-0.87	-0.01	0.2	15-03-21 19:03:03

シナリオ1：学習データ作成、精度測定 その1 コンフィグの作成

The screenshot shows the Predictor web application interface at localhost:8080/predictor/models/3/configs. The page title is "demo (#3)". The navigation bar includes links for "Configs", "Training Data", "Annotations", "Fill Annotation", "Evaluation", and "Inspection". The "Configs" tab is selected and highlighted with a red box.

The main content area displays a table titled "Configs" with two rows:

ID	Config name	Created at	Default config	Retrain	Last retrain status	Duplicate	Description
3	arrow_num_all	15-03-21 18:59:50	★ Default	↻ Retrain	Not Trained	+ Duplicate	
4	nherd_num_all	15-03-21 19:00:21	Set default	↻ Retrain	Not Trained	+ Duplicate	

A blue callout box with the text "Duplicate をクリック" points to the "+ Duplicate" button in the last row of the table. This button is also highlighted with a red box.

Below the table, there is a section titled "Model info" containing the following fields:

Model name (#id)	demo (#3)
Description	(No description)
Model type	multi_class
Owner	Sample User
Tags	

At the bottom, there is a section titled "Download training data" with a file list:

- With IDs
- Status
- Result
- With ID
- With Predictions
- Encoding
- File created at

The file list includes:

- mint_documents-v...tar.gz
- 2015-03-17_1905.zip
- 2015-03-16_1633.zip
- noname
- put

シナリオ1：学習データ作成、精度測定 その1 コンフィグの作成

The screenshot shows a web browser window with the URL `localhost:8080/predictor/models/new`. The page is titled "New model - Sedue Predictor". It contains several input fields and dropdown menus for configuring a new model:

- Name:** default (highlighted with a blue callout bubble containing the text "名前を適当に入力")
- Description:** (empty text area)
- Easy config:**
 - Method:** arow
 - Regularization weight:** 1.0
 - String rule:** space
 - Num rule:** num
- Fields:** Use all fields Select fields
- Batch train config:**
 - Train method:** using_all_data
 - Train times:** 1
- Create model** button (disabled, greyed out)

Sedue Predictor v2.1.0 | © 2014 Preferred Infrastructure

シナリオ1：学習データ作成、精度測定 その1 コンフィグの作成

- コンフィグのパラメータ
 - Name : コンフィグの名前 (なんでも良い)
 - Description : コンフィグの説明 (設定内容のメモ書き)
 - Method : 機械学習のアルゴリズムの選択
 - arow, nherd両方作ってみましょう
 - Regularization weight : 正則化値
 - 例 : 0.01, 0.1, 1, 10, 100など (細かく変えてあまり変わらない)
 - String rule : 文字列からの特徴抽出方法の選択
 - 文書を入れる時にいろんな方法を試します。
 - Number rule : 数値からの特徴抽出方法の選択
 - num,log両方作ってみましょう

シナリオ1：学習データ作成、精度測定 その1 コンフィグの作成

- コンフィグのパラメータ
 - Fields : 学習に使用するフィールドの選択
 - 全部、rawだけ、xだけ、yだけ、zだけなど作ってみましょう
 - Train method : 学習方法の選択
 - 今回はこのままで
 - Train times : 学習回数の選択
 - 何度も学習すると精度が上がる事がある
 - 過学習 (over fit) に注意
 - 学習データに対して過度に最適化されてしまい、実際のデータで上手く行かなくなってしまう

シナリオ1：学習データ作成、精度測定 その1 精度評価の実施

localhost:8080/predictor/models/1/evaluations

sample_model (#1) / Evaluations

Evaluation jobs

	Accuracy	Duration	Finished at
--	----------	----------	-------------

New evaluation

Target config: (All Configs)

Evaluation type: Cross validation

Cross Validation Options

Folds: 5

Training times: 1

Data

All # of training data

100

Submit

mint_documents-v...tar.gz 2015-03-17_1905.zip 2015-03-16_1633.zip noname put すべてを表示

シナリオ1：学習データ作成、精度測定 その1 精度評価は何をしているのか？

- 精度評価（交差検定）とは？
 - 登録してある学習データ（さっき登録したデータです。）の一部（今回は4/5）を使って学習を行ないます。
 - のこりのデータ（今回は1/5）の正解を隠して分類させます。
 - 分類結果と正解を照らし合わせてどれくらい正解できているか確認
 - 上記を分割分（今回は5回）繰り返し、平均を取る。
- テストの種類
 - クローズテスト：学習データにあるデータを使ってテスト
 - オープンテスト：学習データに無いデータを使ってテスト
 - 交差検定はオープンテスト

シナリオ1：学習データ作成、精度測定 その1 精度評価は何をしているのか？

- 精度の考え方
 - 適合率：機械が正解とした内、実際の正解が含まれる割合
 - どれだけ正確に正解を選べるか
 - 再現率：実際の正解の内、機械が当てた割合
 - どれだけ抜け漏れが無いか
 - F値：（適合率 × 再現率 × 2）
 - 適合率、再現率両方を加味した指標

シナリオ1：学習データ作成、精度測定 その1 精度の確認

The screenshot shows the Predictor web application interface at localhost:8080/predictor/models/3/evaluations. The page title is "demo (#3) / Evaluations". The navigation bar includes tabs for Configs, Training Data, Annotations, Fill Annotation, Evaluation (selected), and Inspection.

New evaluation form:

- Target config: `arow_num_all`
- Evaluation type: Cross validation
- Cross Validation Options:
 - Folds: 5
 - Training times: 1
- Data:
 - All (radio button selected)
 - # of training data: 100

Evaluation jobs table:

Status	Detail	Parameters	Accuracy	Duration	Finished at
SUCCESS	View detail...	<ul style="list-style-type: none">Target config : <code>nherd_num_all</code>Evaluation type : Cross validationFolds : 5Training times : 1Data amount : All	50.00%	10s	15-03-21 19:04:08
SUCCESS	View detail...	<ul style="list-style-type: none">Target config : <code>arow_num_all</code>Evaluation type : Cross validationFolds : 5Training times : 1Data amount : All	53.60%	10s	15-03-21 19:04:08

The accuracy value "53.60%" in the second row of the table is highlighted with a red box.

シナリオ1：学習データ作成、精度測定 その1 評価結果詳細の確認

The screenshot shows the Predictor web application interface at localhost:8080/predictor/models/3/evaluations. The page title is "demo (#3) / Evaluations". The navigation bar includes tabs for Configs, Training Data, Annotations, Fill Annotation, Evaluation (which is selected), and Inspection.

New evaluation form:

- Target config: `arow_num_all`
- Evaluation type: Cross validation
- Cross Validation Options:
 - Folds: 5
 - Training times: 1
- Data:
 - All (radio button selected)
 - # of training data: 100

Evaluation jobs table:

Status	Detail	Parameters	Accuracy	Duration	Finished at
SUCCESS	View detail...	<ul style="list-style-type: none">Target config : <code>nherd_num_all</code>Evaluation type : Cross validationFolds : 5Training times : 1Data amount : All	50.00%	10s	15-03-21 19:04:08
SUCCESS	View detail...	<ul style="list-style-type: none">Target config : <code>arow_num_all</code>Evaluation type : Cross validationFolds : 5Training times : 1Data amount : All	53.60%	10s	15-03-21 19:04:08

The "View detail..." link in the second row of the table is highlighted with a red box.

シナリオ1：学習データ作成、精度測定 その1 評価結果詳細 1

The screenshot shows the Predictor web application interface at localhost:8080/predictor/models/3/evaluations/14. The main content area displays the 'Cross validation results' for evaluation #14. A large blue callout box highlights the column headers 'Annotation \ Prediction' and 'Prediction'.

Annotation \ Prediction

	not_work	work	Total	(Recall: σ = 10.53)
not_work	642	359	1001	(64.14 %)
work	570	431	1001	(43.06 %)
Total	(Precision: σ = 0.79) 1212 (52.97 %)	790 (54.56 %)	2002	(Accuracy: 53.60 %)

Prediction

A second blue callout box highlights the row header 'Annotation \ Prediction'.

Evaluation information and parameters

Target config	arow_num_all
Evaluation type	Cross validation
Folds	5
Training times	1
Data amount	All

The browser's address bar shows the URL localhost:8080/predictor/models/3/evaluations/14.

シナリオ1：学習データ作成、精度測定 その1 精度評価詳細の見方

- Confusion Matrixの見方
 - 行 (→) : 機械が分類した内容
 - 列 (↓) : 正解の内容
 - 対角成分 (青) : 正解した件数
 - それ以外 (赤) : 間違えた件数
 - 色の意味: 濃いと多い、薄いと少ない
 - Accuracy: 正確度 (いわゆる精度)

シナリオ1：学習データ作成、精度測定 その1 精度評価詳細2

The screenshot shows a web browser window for the Predictor application at localhost:8080/predictor/models/3/evaluations/15. The page title is "demo (#3) / Evaluations / #15". The navigation bar includes links for Configs, Training Data, Annotations, Fill Annotation, Evaluation (which is selected), and Inspection.

Cross validation results

Result	Annotation \ Prediction	not_work	work	Total	(Recall: σ = 0.0)
Fold #1	not_work	0	1001	1001	(0.00 %)
Fold #2	work	0	1001	1001	(100.00 %)
Fold #3	Total	(Precision: σ = 25.0) 0 (0.00 %)	2002 (50.00 %)	2002	(Accuracy: 50.00 %)
Fold #4					
Fold #5					

Evaluation informations and parameters

Target config	nherd_num_all
Evaluation type	Cross validation
Folds	5
Training times	1
Data amount	All

At the bottom, there is a file list: mint_documents-v...tar.gz, 2015-03-17_1905.zip, 2015-03-16_1633.zip, noname, put, and a "すべてを表示" button.

シナリオ1：学習データ作成、精度測定 その1 まとめ

- シナリオ1の結論
 - これではキーを入力してるかどうかを識別できない。
 - さて、どうするか。。。
- 作戦
 - 収集するデータの量を増やしてみる
 - 学習データを増やしてみる
 - 特徴抽出方法を変えてみる

→ 今回は、収集量を20回に増やしてみる。
意外な方法として文字列として扱ってみる。

シナリオ2：学習データ作成、精度測定 その2 学習データの作成（Edison上の作業）

- 学習データの作成

- センサーの値を20回取得し、その際に、キー入力があったか無かったか正解データを付ける
- PredictorはCSV形式でGUIから学習データを登録できる
- # cd hands_on/edge/session_2
- # node trainer2.js > train2.data
 - このプログラムはCSV形式でデータを作成します。
 - コンソール上でキー入力（強く叩きすぎないで。弱すぎもダメ）
 - 1000個学習データが出来ると終了。
- ヘッター情報を付ける
- # cat header.txt train2.data > train2.csv
- ローカルPCにファイルをコピー

シナリオ2：学習データ作成、精度測定 その2 Predictor上の作業

- モデルの作成
 - new modelボタンをクリックしモデルを作成
 - 一緒に1個コンフィグが作られる
- 学習データの登録
 - training dataを押す
 - upload fileボタンをクリック
 - ファイルを選択、アップロード

シナリオ2：学習データ作成、精度測定 その2 Predictor上の作業

- コンフィグの作成
 - 色々なコンフィグを作成してみよう
- 精度評価
 - Evaluationを選択
 - target configでall configを選択、他はデフォルトでOK
 - submitボタンを押す
 - Accuracyを確認
- 分類理由の確認
 - Configを選択
 - retrainを実施
 - Training Dataを確認

シナリオ2：学習データ作成、精度測定 その2 評価結果の確認

The screenshot shows a web browser window for the Predictor application at localhost:8080/predictor/models/2/evaluations. The page has a blue header with the title "jaws_demo (#2) / Evaluations". Below the header, there is a navigation bar with tabs: Configs, Training Data, Annotations, Fill Annotation, Evaluation (which is selected), and Inspection.

New evaluation

Target config: default

Evaluation type: Cross validation

Cross Validation Options

Folds: 5

Training times: 1

Data

All

of training data

100

Evaluation jobs

Status	Detail	Parameters	Accuracy	Duration	Finished at
SUCCESS	View detail...	<ul style="list-style-type: none">Target config : defaultEvaluation type : Cross validationFolds : 5Training times : 1Data amount : All	81.20%	10s	15-03-20 18:00:24

At the bottom of the browser window, there is a file list:

- mint_documents-v...tar.gz
- 2015-03-17_1905.zip
- 2015-03-16_1633.zip
- noname
- put

シナリオ2：学習データ作成、精度測定 その2 評価結果の詳細

The screenshot shows a web browser displaying the Predictor application at localhost:8080/predictor/models/2/evaluations/9. The page title is "jaws_demo (#2) / Evaluations / #9". The navigation bar includes links for "Configs", "Training Data", "Annotations", "Fill Annotation", "Evaluation" (which is selected), and "Inspection".

Cross validation results

Result	Annotation \ Prediction		rest	work	Total	(Recall: $\sigma = 4.39$)
Fold #1	rest		428	72	500	(85.60 %)
Fold #2	work		116	384	500	(76.80 %)
Fold #3	Total	(Precision: $\sigma = 2.76$)	544 (78.68 %)	456 (84.21 %)	1000	(Accuracy: 81.20 %)
Fold #4						
Fold #5						

Evaluation informations and parameters

Target config	default
Evaluation type	Cross validation
Folds	5
Training times	1
Data amount	All

The bottom of the browser window shows the OS X Dock with several open files and folders.

シナリオ2：学習データ作成、精度測定 その2

- 学習の実施
 - Configを選択
 - retrainを実施
- 分類理由の確認
 - Training Dataを確認
 - Inspectionの実施

シナリオ2：学習データ作成、精度測定 その2 学習の実施

The screenshot shows the Predictor web application interface at localhost:8080/predictor/models/1/configs. The page title is "sample_model (#1)". The navigation bar includes links for Predictor, Models, Jobs, and Files, with "Sample User" selected. The main content area has tabs for Configs, Training Data, Annotations, Fill Annotation, Evaluation, and Inspection, with "Configs" selected. A red box highlights the "Configs" tab. Below it, a table lists one configuration entry:

ID	Config name	Created at	Default config	Retrain	Last retrain status	Duplicate	Description
1	default	15-03-10 03:25:20	★ Default	Retrain	success 15-03-19 10:11:53	Duplicate	Auto-generated co...

A blue callout box with white text points to the "Retrain" button in the table row, containing the text "retrainボタンをクリック".

Below the table, there is a section titled "Model info" with the following details:

Model name (#id)	sample_model (#1)
Description	Auto-generated model sample.
Model type	multi_class
Owner	Sample User
Tags	auto_generated

At the bottom, there is a section titled "Download training data" with a table header:

<input type="checkbox"/> With IDs	Status	Result	With ID	With Predictions	Encoding	File created at
-----------------------------------	--------	--------	---------	------------------	----------	-----------------

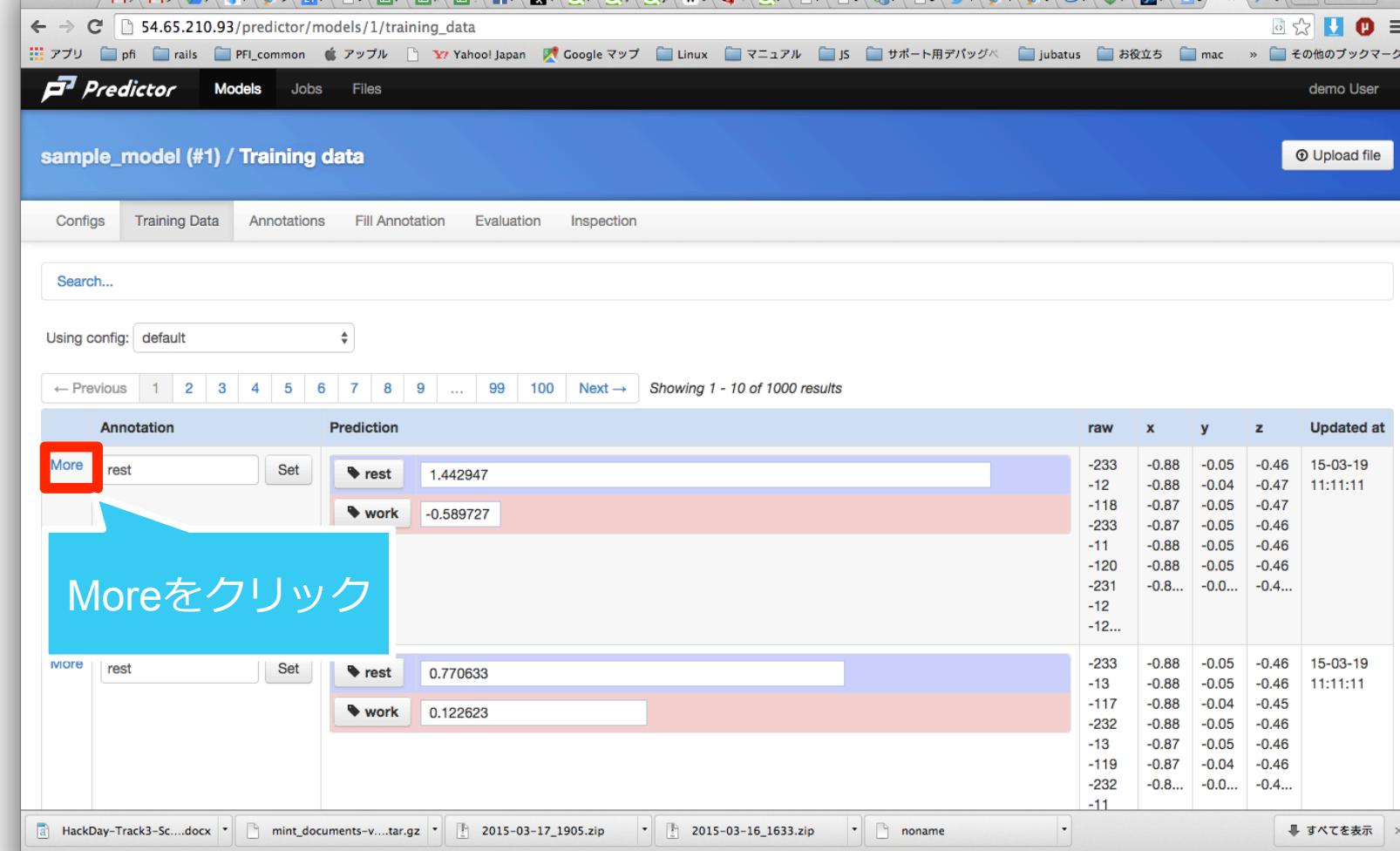
The status bar at the bottom shows several file icons: HackDay-Track3-Sc...docx, mint_documents-v...tar.gz, 2015-03-17_1905.zip, 2015-03-16_1633.zip, and noname, along with a "すべてを表示" (Show all) button.

シナリオ2：学習データ作成、精度測定 その2 分類理由の確認

The screenshot shows a web-based application for managing machine learning models. The URL is 54.65.210.93/predictor/models/1/training_data. The interface includes a navigation bar with links like 'Configs', 'Training Data', 'Annotations', 'Fill Annotation', 'Evaluation', and 'Inspection'. The main area displays 'sample_model (#1) / Training data' with a search bar and a table of results. The table has columns: Annotation, Prediction, raw, x, y, z, and Updated at. A red box highlights the first row, which shows 'rest' as the annotation, '1.442947' as the prediction, and a list of raw data points. A large blue callout bubble with the text '分類結果を表示' (Classification results displayed) points to this row. The bottom of the screen shows a toolbar with various file icons.

Annotation	Prediction	raw	x	y	z	Updated at
More rest	1.442947	233 -12 -118 233 -11 -120 -231 -12 -12...	-0.88 -0.88 -0.87 -0.87 -0.88 -0.88 -0.8...	-0.05 -0.04 -0.05 -0.05 -0.05 -0.05 -0.0...	-0.46 -0.47 -0.47 -0.46 -0.46 -0.46 -0.4...	15-03-19 11:11:11
More rest	-0.589727	-233 -13 -117 -232 -13 -119 -232 -11	-0.88 -0.88 -0.88 -0.88 -0.87 -0.87 -0.8...	-0.05 -0.05 -0.04 -0.05 -0.05 -0.04 -0.0...	-0.46 -0.46 -0.45 -0.46 -0.46 -0.46 -0.4...	15-03-19 11:11:11

シナリオ2：学習データ作成、精度測定 その2 分類理由の確認



The screenshot shows a web-based application for managing machine learning models. The URL is 54.65.210.93/predictor/models/1/training_data. The interface includes a navigation bar with links to 'Models', 'Jobs', and 'Files'. The main area is titled 'sample_model (#1) / Training data' and contains tabs for 'Configs', 'Training Data', 'Annotations', 'Fill Annotation', 'Evaluation', and 'Inspection'. The 'Training Data' tab is selected.

The 'Training Data' section displays 1000 results. It includes a search bar, a dropdown for 'Using config: default', and a navigation bar with pages 1 through 100. Below these are two rows of annotations:

Annotation	Prediction	raw	x	y	z	Updated at
More rest	1.442947	-233 -12 -118 -233 -11 -120 -231 -12 -12...	-0.88 -0.88 -0.87 -0.87 -0.88 -0.88 -0.8...	-0.05 -0.04 -0.05 -0.05 -0.05 -0.05 -0.0...	-0.46 -0.47 -0.47 -0.46 -0.46 -0.46 -0.4...	15-03-19 11:11:11
More work	-0.589727					
More rest	0.770633	-233 -13 -117 -232 -13 -119 -232 -11	-0.88 -0.88 -0.88 -0.88 -0.87 -0.87 -0.8...	-0.05 -0.05 -0.04 -0.05 -0.05 -0.04 -0.0...	-0.46 -0.46 -0.45 -0.46 -0.46 -0.46 -0.4...	15-03-19 11:11:11
More work	0.122623					

A large blue callout box with the text 'Moreをクリック' (Click More) is overlaid on the first row of annotations. The 'More' button in the first row is highlighted with a red box.

シナリオ2：学習データ作成、精度測定 その2 分類理由の確認

シナリオ2：学習データ作成、精度測定 その2 分類理由の確認

赤が分類にNegativeに効いている特徴

x	-0.88	-0.88	-0.87	-0.87	-0.88	-0.88	-0.87	-0.88	-0.88	-0.87	-0.87	-0.88	-0.88	-0.88	-0.88	-0.88	-0.88	-0.88	-0.88	-0.88	-0.88	-0.88	-0.88	-0.88		
z	-0.46	-0.47	-0.47	-0.46	-0.46	-0.46	-0.46	-0.46	-0.47	-0.47	-0.47	-0.46	-0.46	-0.46	-0.46	-0.46	-0.46	-0.46	-0.46	-0.46	-0.46	-0.46	-0.46	-0.46	-0.47	
y	-0.05	-0.04	-0.05	-0.05	-0.05	-0.05	-0.04	-0.05	-0.05	-0.05	-0.05	-0.05	-0.05	-0.05	-0.04	-0.05	-0.05	-0.04	-0.05							
	-0.05	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.05	-0.05	-0.05	-0.04	-0.05														

work

Field	Content
raw	-233 -12 -118 -233 -11 -120 -231 -12 -120 -231 -13 -118 -232 -12 -117 -232 -12 -117 -231 -11 -118 -232 -12 -119 -232 -13 -120 -231 -13 -121 -231 -12 -117 -232 -12 -118 -233 -12 -119 -232 -11 -119 -232 -11 -119 -232 -11 -119 -232 -11 -119 -232 -11 -121 -232 -11 -120 -232 -11 -119 -232 -11 -117 -233 -12 -117 -232 -13 -118 -232 -12 -119 -233 -11 -119 -232 -12 -120 -231 -12 -119 -233 -11 -119 -232 -13 -117 -231 -13 -118 -232 -12 -119
x	-0.88 -0.88 -0.87 -0.87 -0.88 -0.88 -0.87 -0.87 -0.88 -0.88 -0.87 -0.87 -0.88 -0.88 -0.88 -0.88 -0.88 -0.88 -0.88 -0.88 -0.88 -0.88 -0.88 -0.88 -0.88 -0.88
z	-0.46 -0.47 -0.47 -0.46 -0.46 -0.46 -0.46 -0.46 -0.47 -0.47 -0.47 -0.46 -0.46 -0.46 -0.46 -0.46 -0.46 -0.46 -0.46 -0.46 -0.46 -0.46 -0.46 -0.46 -0.46 -0.46 -0.47
v	0.05 -0.04 0.05 0.05 0.05 -0.04 0.05

シナリオ2：学習データ作成、精度測定 その2 分類理由の確認

The screenshot shows a web browser window with the URL 54.65.210.93/predictor/models/1/inspections. The page title is "sample_model (#1) / Inspections". The navigation bar includes tabs for "Configs", "Training Data", "Annotations", "Fill Annotation", "Evaluation", and "Inspection". The "Inspection" tab is highlighted with a red box. On the left, there's a form for "New inspection" with fields for "Target config" (set to "default"), "Inspection type" (set to "Feature ranking"), and "Sample size" (set to "30"). Below the form is a "Submit" button, also highlighted with a red box. To the right, there's a table titled "Inspection jobs" with columns: Status, Detail, Parameters, Duration, and Finished at. The table currently has no data. At the bottom of the page, there's a file list: "HackDay-Track3-Sc...docx", "mint_documents-v...tar.gz", "2015-03-17_1905.zip", "2015-03-16_1633.zip", and "noname". A blue callout box with white text "Submitボタンをクリック" (Click the Submit button) points to the "Submit" button.

シナリオ2：学習データ作成、精度測定 その2 分類理由の確認

The screenshot shows the Predictor web application interface at localhost:8080/predictor/models/2/inspections. The page title is "jaws_demo (#2) / Inspections". The navigation bar includes links for "Configs", "Training Data", "Annotations", "Fill Annotation", "Evaluation", and "Inspection". The "Inspection" tab is active.

New inspection

- Target config: default
- Inspection type: Feature ranking
- Sample size: 30 (between 5 and 30)

Submit

Inspection jobs

Status	Detail	Parameters	Duration	Finished at
SUCCESS	View detail...	<ul style="list-style-type: none">Target config : defaultInspection type : Feature rankingSample size : 30	3s	15-03-21 22:04:22

A large blue callout box with the text "詳細を確認" (Check details) points to the "View detail..." link in the table row.

At the bottom, there is a file list:

- HackDay-Track3-Sc...docx
- mint_documents-v...tar.gz
- 2015-03-17_1905.zip
- 2015-03-16_1633.zip
- noname

On the right, there is a "すべてを表示" (Show all) button and a close button (X).

シナリオ2：学習データ作成、精度測定 その2 分類理由の確認

The screenshot shows a web browser displaying the Predictor application at localhost:8080/predictor/models/2/inspections/19. The page title is "jaws_demo (#2) / Inspections / #19". The navigation bar includes links for "Configs", "Training Data", "Annotations", "Fill Annotation", "Evaluation", and "Inspection". The "Inspection" tab is active. Below the tabs, there are buttons for "Summary", "rest", and "work". A large blue callout box contains the text "効いている特徴のランキングを表示". The main content area is titled "Summary" and displays a table comparing feature values for "rest" and "work" categories.

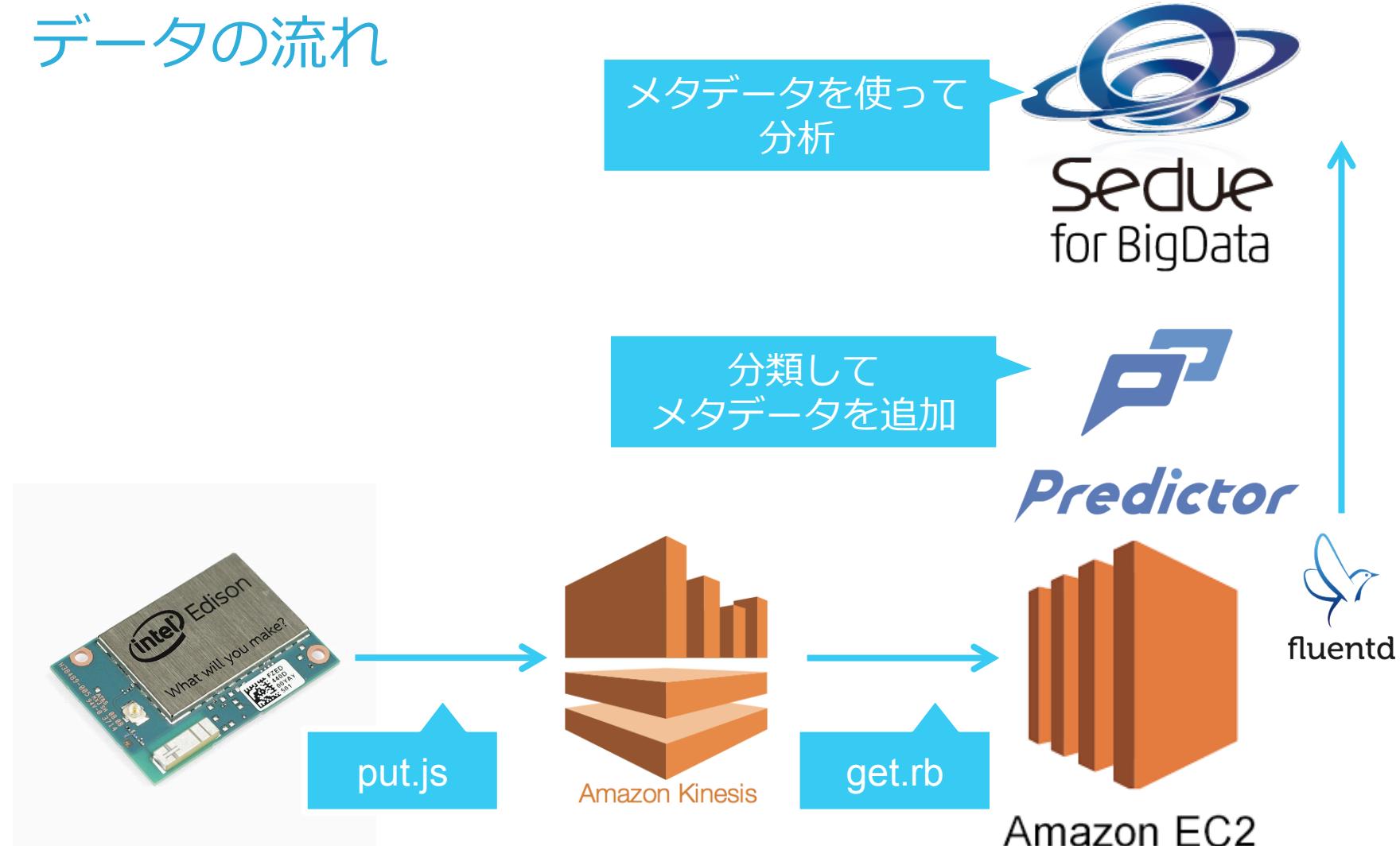
Field	Feature	▼ rest	▼ work
raw	-242	0.231528	-0.231528
raw	-17	0.208109	-0.200911
raw	-19	0.207573	-0.207573
raw	-107	0.192048	-0.192048
y	0.02	0.180156	-0.180156
raw	-247	0.172011	-0.172011
x	-0.93	0.159707	-0.159707
raw	-244	0.148293	-0.148293
raw	-3	0.135637	-0.135637
raw	5	0.135349	-0.135349
raw	-238	0.131721	-0.131721
raw	-228	0.130620	-0.130620
raw	-235	0.128165	-0.128165
raw	-111	0.127487	-0.127487

シナリオ2：学習データ作成、精度測定 その2 まとめ

- シナリオ2の結論
 - 精度を出すためには色々な工夫が必要
 - データの取り方、特徴の取り方、パラメータを変えてみたり
 - 普通はプログラムを書き直して行なう
 - 工夫をするたびに精度測定が必要
 - 普通は上記の組み合わせごとに学習させて分類して集計して。。。
 - 分類した理由を確認
 - 普通は出来ない。（モデルをダンプして解析する必要がある）
- Sedue Predictorは上記の作業を協力に支援！（宣伝）

では実際に分析をしてみましょう！

シナリオ3：実際に分析を実施してみる データの流れ



シナリオ3：実際に分析を実施してみる

- Sedue Predictorの準備
 - sample_modelを選択
 - defaultコンフィグを修正する
 - シナリオ2で精度の良かったコンフィグと同じ設定にする
 - シナリオ2で作成した学習データを登録
 - コンフィグをretrainする

今回はsample modelを使って分類します。

(事前にしてある設定を使うため)

シナリオ3：実際に分析を実施してみる

The screenshot shows a web browser window for the Predictor application at localhost:8080/predictor/models/1/configs. The page title is "sample_model (#1)". The main content area has tabs for "Configs", "Training Data", "Annotations", "Fill Annotation", "Evaluation", and "Inspection". The "Configs" tab is selected, displaying a table with one row:

ID	Config name	Created at	Default config	Retrain	Last retrain status	Duplicate	Description
1	default	15-03-19 25:20	★ Default	Retrain	success 15-03-19 10:11:53	+ Duplicate	Auto-generated co...

Below the table is a section titled "Model info" with the following details:

Model name (#id)	sample_model
Description	Auto-generated
Model type	multi_class
Owner	Sample User
Tags	auto_generated

A large blue arrow points downwards from the "Model info" section towards the "Download training data" section. A blue callout bubble contains the Japanese text "下にスクロールしてください" (Please scroll down). The "Download training data" section includes a checkbox for "With IDs" and a dropdown menu for "With predictions using config". The file list at the bottom shows several files: "HackDay-Track3-Sc...docx", "mint_documents-v...tar.gz", "2015-03-17_1905.zip", "2015-03-16_1633.zip", and "noname".

シナリオ3：実際に分析を実施してみる

54.64.224.68/predictor/models/1/configs

SJIS

Execute

Dump state of the model and config

Status	Result	Target config	Encoding	File created at
default				

Shows up to 5 latest executions ([view all files](#))

Encoding

SJIS

Execute

Clear training data

CLEAR TRAINING DATA

学習データをクリア

Delete everything in this model and the moae!

DELETE EVERYTHING IN THIS MODEL AND THE MODEL

Sedue Predictor v2.1.0 | © 2014 Preferred Infrastructure

HackDay-Track3-Sc...docx mint_documents-v...tar.gz 2015-03-17_1905.zip 2015-03-16_1633.zip noname すべてを表示

シナリオ3：実際に分析を実施してみる

The screenshot shows a web browser window for the Predictor application at localhost:8080/predictor/models/1/configs. The page title is "sample_model (#1)". The top navigation bar includes links for Predictor, Models, Jobs, and Files, and a user account for "Sample User". Below the title, there's a button labeled "Edit model". The main content area has tabs for Configs, Training Data, Annotations, Fill Annotation, Evaluation, and Inspection. The "Configs" tab is selected, displaying a table with one row:

ID	Config name	Created at	Default config	Retrain	Last retrain status	Duplicate	Description
1	default			<input type="button" value="Retrain"/>	success 15-03-19 10:11:53	<input type="button" value="Duplicate"/>	Auto-generated co...

A red box highlights the "default" link in the table. A blue callout box with the text "defaultを選択" (Select default) points to this link. Below the table, the "Model info" section contains the following details:

Model name (#id)	sample_model (#1)
Description	Auto-generated model sample.
Model type	multi_class
Owner	Sample User
Tags	auto_generated

At the bottom, there's a section titled "Download training data" with a table header:

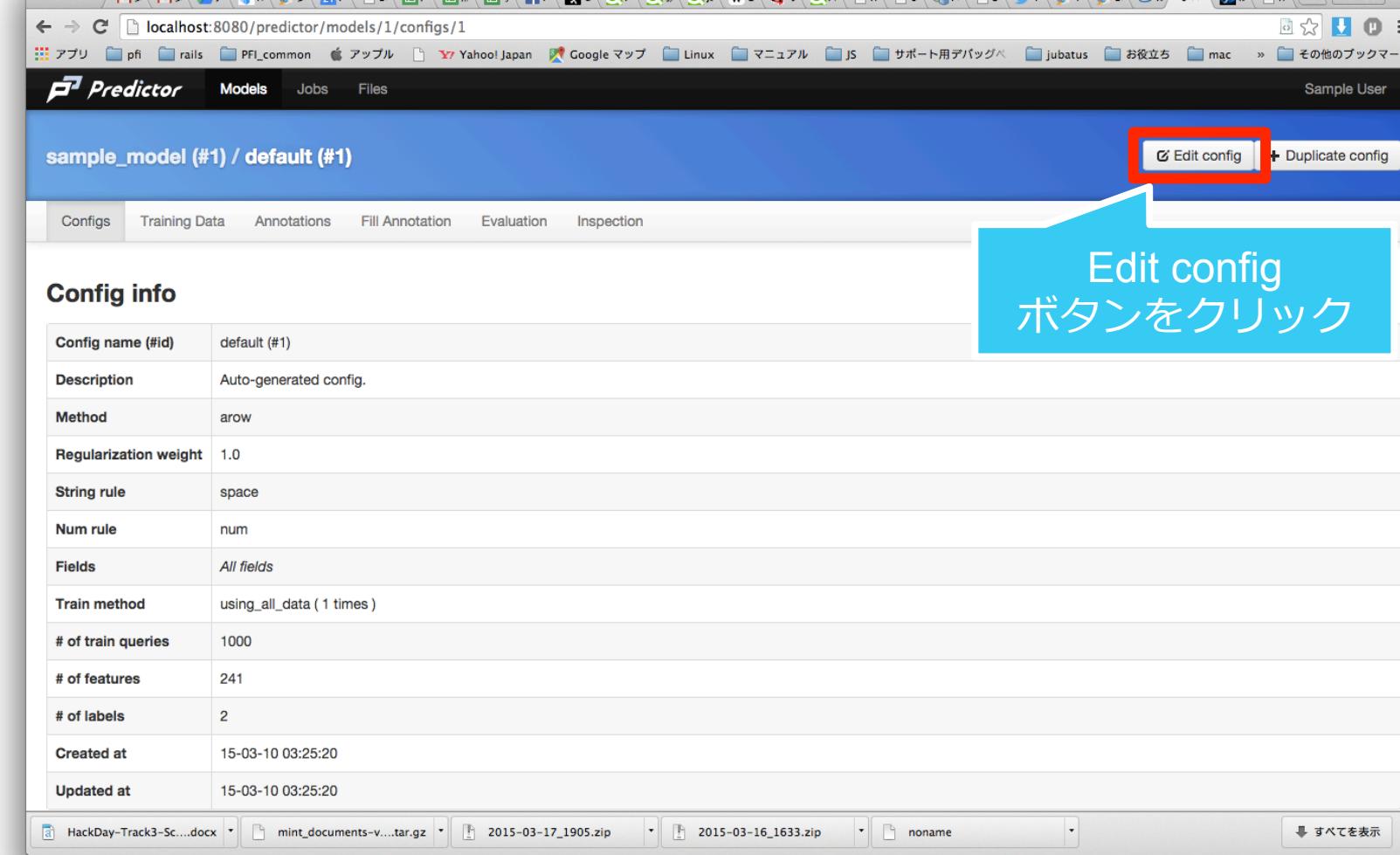
<input type="checkbox"/> With IDs	Status	Result	With ID	With Predictions	Encoding	File created at
-----------------------------------	--------	--------	---------	------------------	----------	-----------------

The table body shows several file entries:

- HackDay-Track3-Sc...docx
- mint_documents-v...tar.gz
- 2015-03-17_1905.zip
- 2015-03-16_1633.zip
- noname

On the right side of the download table, there are buttons for "すべてを表示" (Show all) and "x".

シナリオ3：実際に分析を実施してみる



The screenshot shows a web browser window for the Predictor application at localhost:8080/predictor/models/1/configs/1. The page title is "sample_model (#1) / default (#1)". At the top right, there are two buttons: "Edit config" (highlighted with a red box) and "Duplicate config". Below the title, there is a navigation bar with tabs: "Configs", "Training Data", "Annotations", "Fill Annotation", "Evaluation", and "Inspection". The main content area is titled "Config info" and contains a table with the following data:

Config name (#id)	default (#1)
Description	Auto-generated config.
Method	arow
Regularization weight	1.0
String rule	space
Num rule	num
Fields	All fields
Train method	using_all_data (1 times)
# of train queries	1000
# of features	241
# of labels	2
Created at	15-03-10 03:25:20
Updated at	15-03-10 03:25:20

At the bottom of the page, there is a toolbar with several file icons and a "すべてを表示" (Show all) button. A large blue callout box on the right side of the page points to the "Edit config" button with the text "Edit config ボタンをクリック" (Click the Edit config button).

シナリオ3：実際に分析を実施してみる

The screenshot shows a web application interface for editing a configuration. The URL is `localhost:8080/predictor/models/1/configs/1/edit`. The page has a navigation bar with tabs: Configs, Training Data, Annotations, Fill Annotation, Evaluation, and Inspection. The 'Configs' tab is selected.

The main form contains the following fields:

- Name: default
- Description: Auto-generated config.
- Easy config:
 - Method: arow
 - Regularization weight: 1.0
 - String rule: space
 - Num rule: num
- Fields:
 - Use all fields
 - Select fields
- Batch train config:
 - Train method: using_all_data
 - Train times: 1

A blue callout box with a white arrow points from the right towards the configuration fields, containing the text: "シナリオ2でベストだった
コンフィグを設定".

At the bottom center of the form is a red-bordered button labeled "Update config".

The browser's address bar shows the current URL. Below the browser window is a Mac OS X dock with several open application icons.

シナリオ3：実際に分析を実施してみる

localhost:8080/predictor/models/1/configs

Predictor Models Jobs Files Sample User

sample_model (#1) Edit model

Configs Training Data Annotations Fill Annotation Evaluation Inspection

Configs

ID	Config name	Created at	Default config	Retrain	Last retrain status	Duplicate	Description
1	default	15-03-10 03:25:20	★ Default	 Retrain	success 15-03-19 10:11:53	 Duplicate	Auto-generated co...

Model info

Model name (#id)	sample_model (#1)
Description	Auto-generated model sample.
Model type	multi_class
Owner	Sample User
Tags	auto_generated

Download training data

<input type="checkbox"/> With IDs	Status	Result	With ID	With Predictions	Encoding	File created at
With predictions using config						

HackDay-Track3-Sc...docx mint_documents-v...tar.gz 2015-03-17_1905.zip 2015-03-16_1633.zip noname すべてを表示 x

シナリオ3：実際に分析を実施してみる データをkinesisに送る（Edison上の作業）

- プログラムを修正する

- session_3/put.jsを修正

```
var cognitoParams = {  
    AccountId: "ここをうめる",  
    RoleArn: "arn:aws:iam:: * * ", //ここうめる  
    IdentityPoolId: "us-east-1: * * " //ここうめる  
};
```

...

```
var params = { Data: JSON.stringify(data), PartitionKey: "demo",  
    StreamName: "ここを埋める" };
```

- プログラムを実行する

- # node put.js

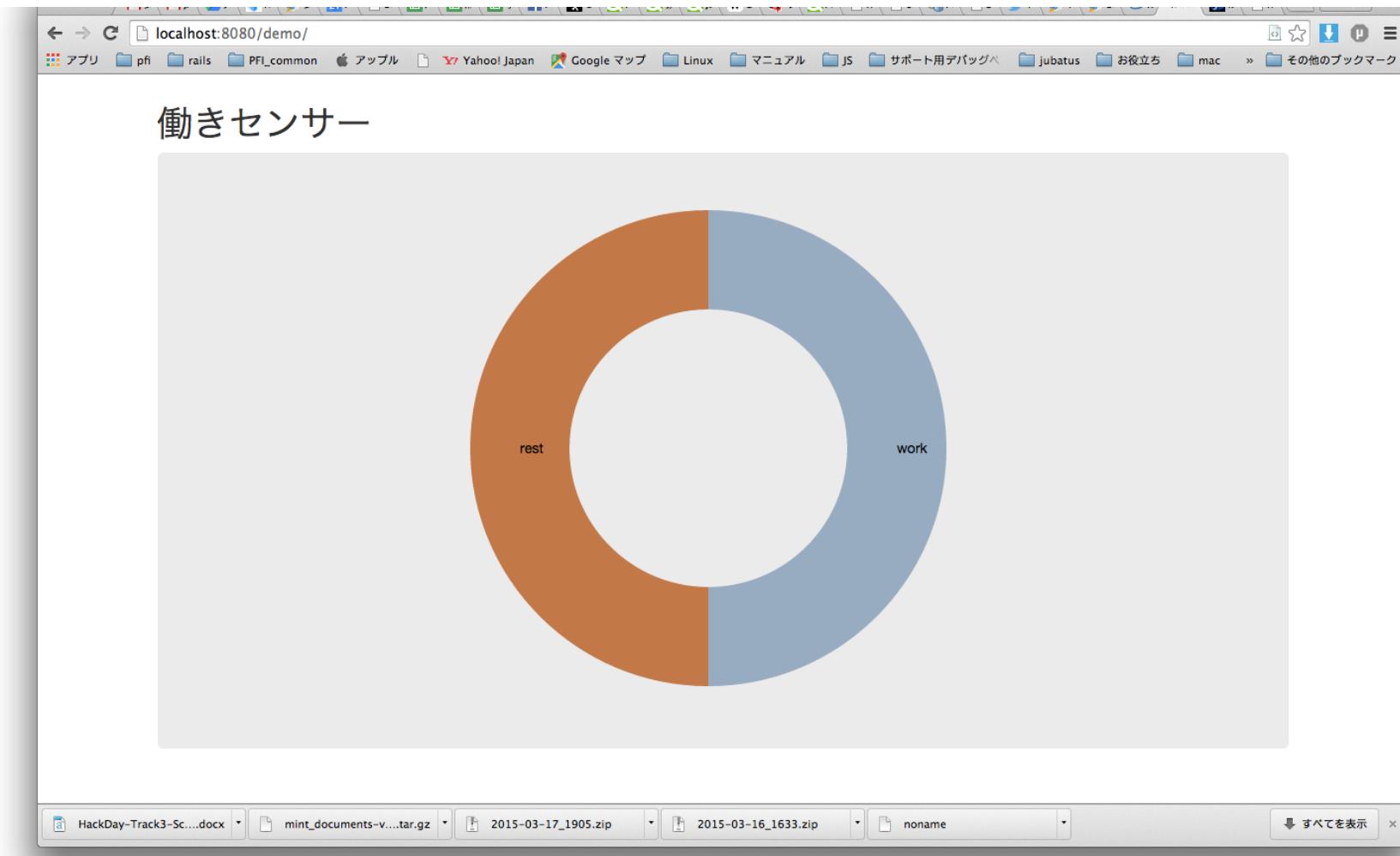
シナリオ3：実際に分析を実施してみる データを受け取る（EC2上の作業）

- EC2にログイン
 - \$ ssh demo@IPアドレス
 - パスワード : jaws2015
- プログラムを実行
 - \$ ruby get.rb --stream <stream名> --tag sfbd.query.add

シナリオ3：実際に分析を実施してみる 確認する

- 登録結果を確認する
 - ブラウザで確認
 - <http://ホスト名/demo>

シナリオ3：実際に分析を実施してみる 結果の確認



シナリオ3：実際に分析を実施してみる 結果の確認



まとめ

- 今日は機械学習を使ってデータの分析をしてみました。
- 機械学習の使い方の流れ分かりましたか？
 - データを収集する
 - 学習データを作成する
 - 特徴を抽出する
 - 学習をする
 - 精度評価を行なう
- センサーデータはただ収集して可視化するだけでは意味がありません。機械学習を使って分析しましょう！

最後に

- 環境を削除するのを忘れずに。
- ご質問とか有ればお気軽に！
- Sedue製品にご興味のある方はぜひお声がけください！
 - 体験版VMなど、ご用意しております！



Copyright © 2006-2015

Preferred Infrastructure All Right Reserved.