意思決定のための ビジュアルプログラミングによる データ分析を支援する Webアプリケーションの開発

Developing Web Applications to Support Data Analysis by Visual Programming for Decision Making

Keniti Numata
u055017@st.pu-toyama.ac.jp

Graduate School of Engineering
Department of Information Systems Engineering
Toyama Prefectural University

Teams, 9:50-10:15 Friday., December 4, 2020.

1.1 本研究の背景

背景

近年，企業などではSociety5.0に向けた世間に溢れる様々な情報を収集し，ビックデータとして様々な処理や分析によって情報を扱うことが増えている．これらのデータは，膨大であるため一般的にプログラミング言語を使って処理される．

ビッグデータについて

1. ビッグデータとは，大量で高頻度な多様性があるデータとして定義される．
2. ビッグデータの例として，ライフログデータや顧客データ，購買データなどがある．
3. 需要の予測やコストの削減，仕入れなどの最適化など様々な活用方法がある．

1.2 本研究の目的

現在の問題点

1. ユーザビリティーの低さ
   プログラミング初心者にとって，プログラミングは手にかけづらい．
2. データの取り扱い
   データを簡単に分析する手法が存在しない．

本研究の目的

・プログラミング初心者でも扱いやすい環境の開発
・データを処理できるように開発
・外部に公開し，複数人から利用できるようにする．

2.1 Society 5.0 と web サービス

Society 5.0

Society 5.0 とは，サイバー空間（仮想空間）とフィジカル空間（現実空間）を高度に融合させたシステムにより、経済発展と社会的課題の解決を両立する人間中心の社会のことである．

Society 5.0 と web サービス

サイバー空間上で web サービスとして Blockly を提供し，ユーザーがフィジカル空間上でデータを分析できるようにすることで Society 5.0 を実現する．
また，このような IT を利用した人々の生活をあらゆる面でより良い方向に変化させるようなことをデジタルトランスフォーメーションという．

2.2 サーバサイドプログラミング

サーバサイドプログラミング

本研究では，HTTP 通信を行ったあとにユーザーからの入力に対しての処理を，非同期通信を使って Web サーバ上でプログラムの実行が要求され，結果をウェブブラウザに対して送信するシステムを開発する．
クライアントサイドコードに使うことができる言語は，HTML,CSS,Javascript であるが，サーバーサイドによる処理を挟むことでそれ以外の言語 (Perl,PHP,Python,Ruby など) を使う事ができるようになる．

ビジュアルプログラミング言語

プログラムをテキストで記述するのではなく，視覚的なオブジェクトで記述するプログラミング言語のこと．視覚的でわかりやすいものが多いため，プログラムの組み立て方を学ぶのに有効であると注目されている．

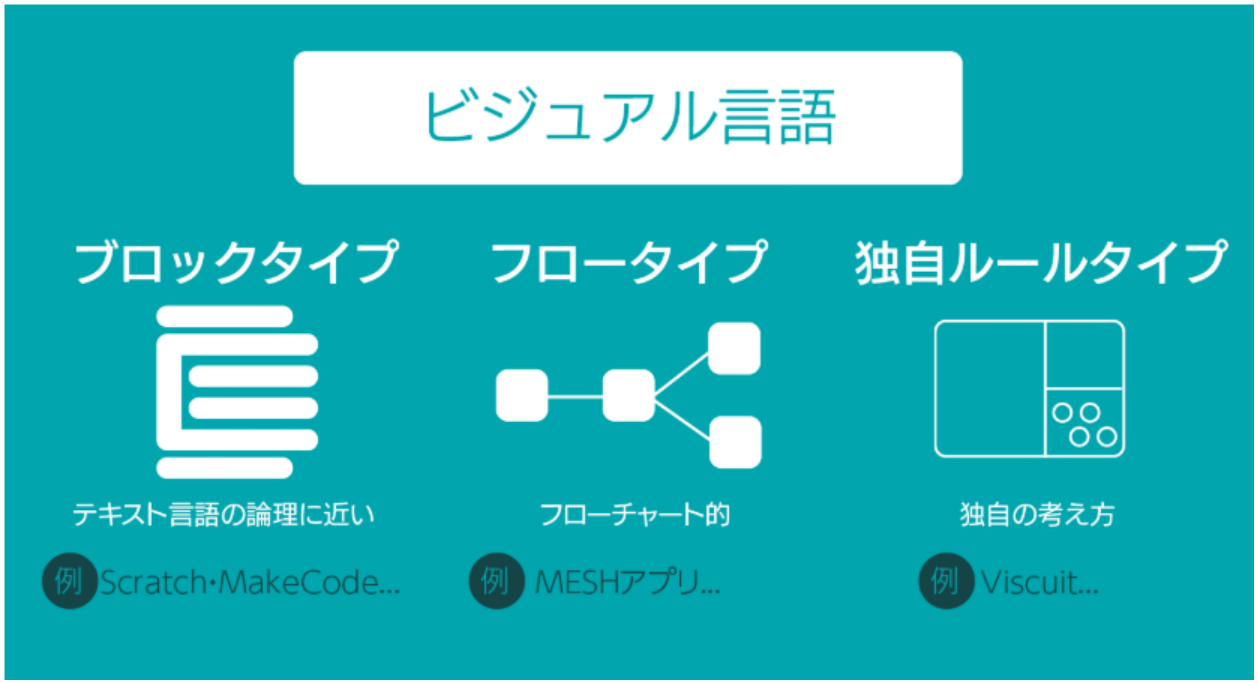

Figure 1: ビジュアルプログラミング

3.2 ブロックタイプのビジュアルプログラミング言語

ブロックタイプのビジュアルプログラミング言語

機械学習（人工知能・AI）を使って課題を解決するクラウドサービスの MAGELLAN BLOCKS（BLOCKS）や教育用作られ様々なアプリケーションに応用して使われている Blockly などがある．
応用して使われているサービスとして Scratch や MakeCode が存在する．

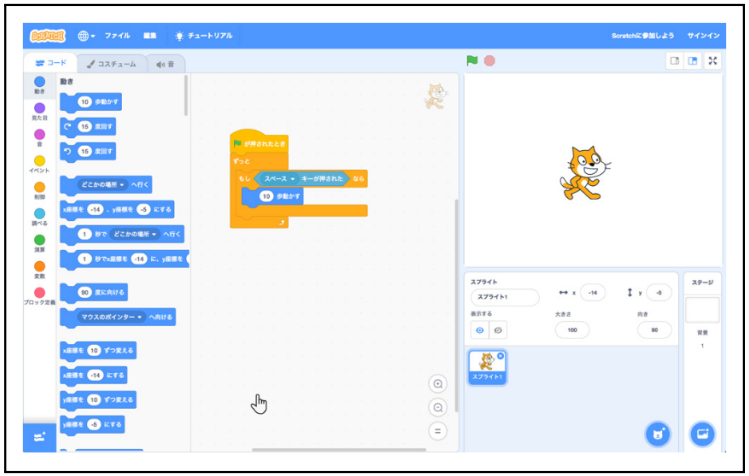

Figure 2: Scratch

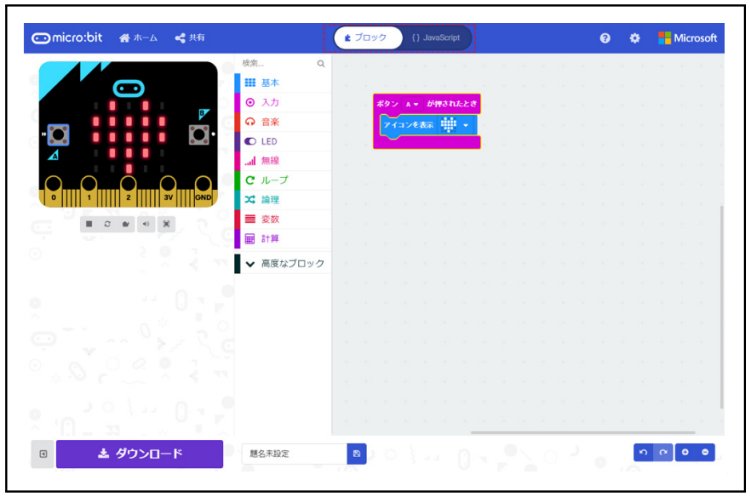

Figure 3: MakeCode

3.3 ビジュアルプログラミング言語 (Blockly)

Blockly

Google が提供しているビジュアルプログラミング言語のライブラリ．簡単な記述で自分だけのビジュアルプログラミング言語を作ることができる．
また，カスタムブロックという機能があり，もともとあるブロックの他にユーザが好きなブロックを作成することができる．

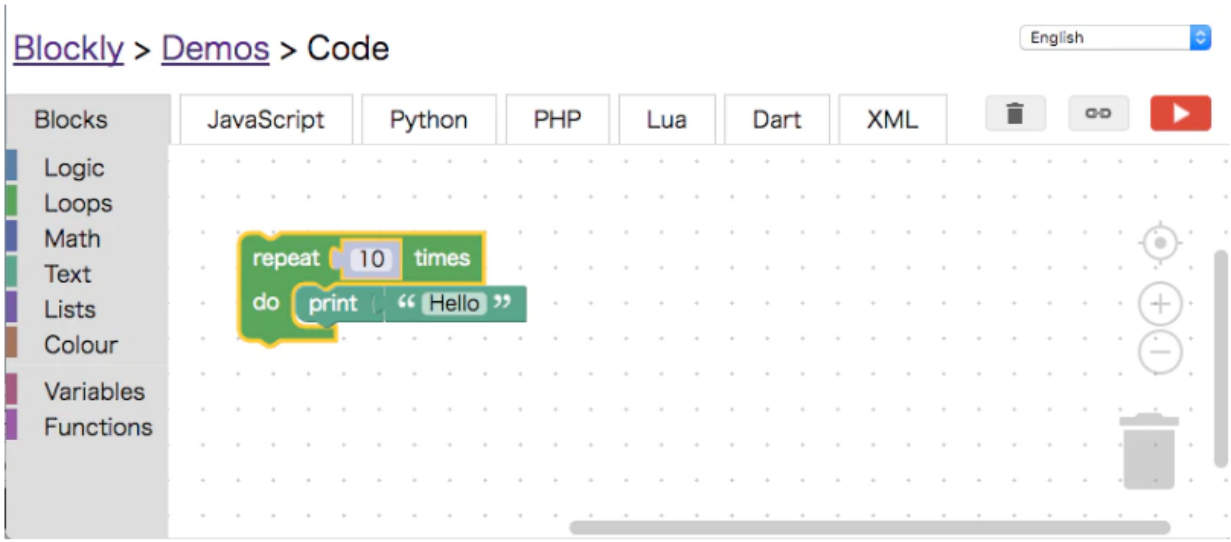

Figure 4: Blockly

3.4.1 ブロックの定義

作成したいブロックの外観とブロックに接続する数値やテキストをここで定義する．
外観は，ブロックの色やブロックの接続 (構文ブロックと値ブロック)，表示する文字等がある．また，ブロック内の空きに何を入力 (input) として何を出力 (output) とするかなど決める．

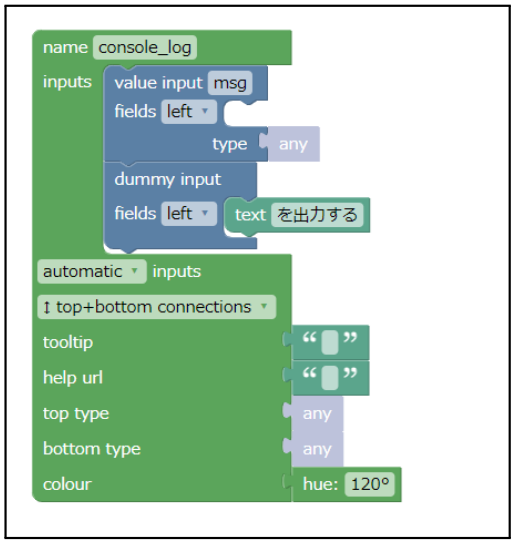

Figure 5: console log に結果を出力するブロックの定義

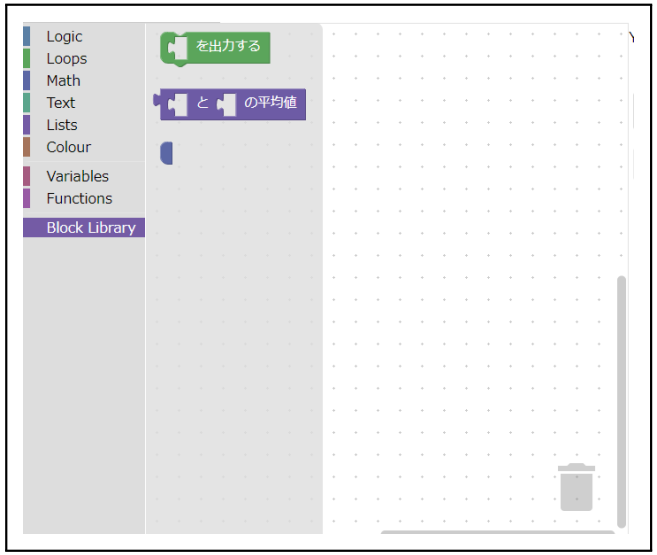

Figure 6: toolbox

3.4.2 コードの生成

コードの生成では，ブロックの動作の定義を行う．例えば，平均値を出すブロックを作成するときは，平均を出す計算部分を動作の定義で書く．

erage.js > ...

```
Blockly.JavaScript['average'] = function(block) {
  var value_v1 = Blockly.JavaScript.valueToCode(block, 'v1', Blockly.JavaScript.ORDER_ATOMIC);
  var value_v2 = Blockly.JavaScript.valueToCode(block, 'v2', Blockly.JavaScript.ORDER_ATOMIC);
  // TODO: Assemble JavaScript into code variable.
  var code = '(' + value_v1 + '+' + value_v2 + ')/2';
  // TODO: Change ORDER_NONE to the correct strength.
  return [code, Blockly.JavaScript.ORDER_NONE];
};
```

Figure 7: 動作の定義

3.4.3 ブロックのカテゴリーと配置決め

作ったブロックをどこのカテゴリーに入れるかを決める．

4.1 提案手法

従来の Blockly の問題点

- データを分析するとき，外部からデータを読み込む必要があるが Blockly のデモコードには外部からのファイルの読み込みが実装されていない点.
- Blockly はブラウザをベースにしてクライアント側で javascript によって処理を行うサーバ・サイド・インクルード (SSI) であるため，データ分析関連の処理をできるライブラリを使うことができない点.

提案手法

- Asynchronous JavascriptAnd XML(Ajax) による通信
- 機械学習ライブラリの導入
- CSV ファイルの読み込みブロックの作成

Ajax 通信とは

通常，HTTP とは Web ページなどをやり取りする技術（プロトコル）で，ブラウザとサーバーとの間で使われている．Ajax を使うと，ブラウザに代わって JavaScript がサーバーと HTTP 通信することができる．

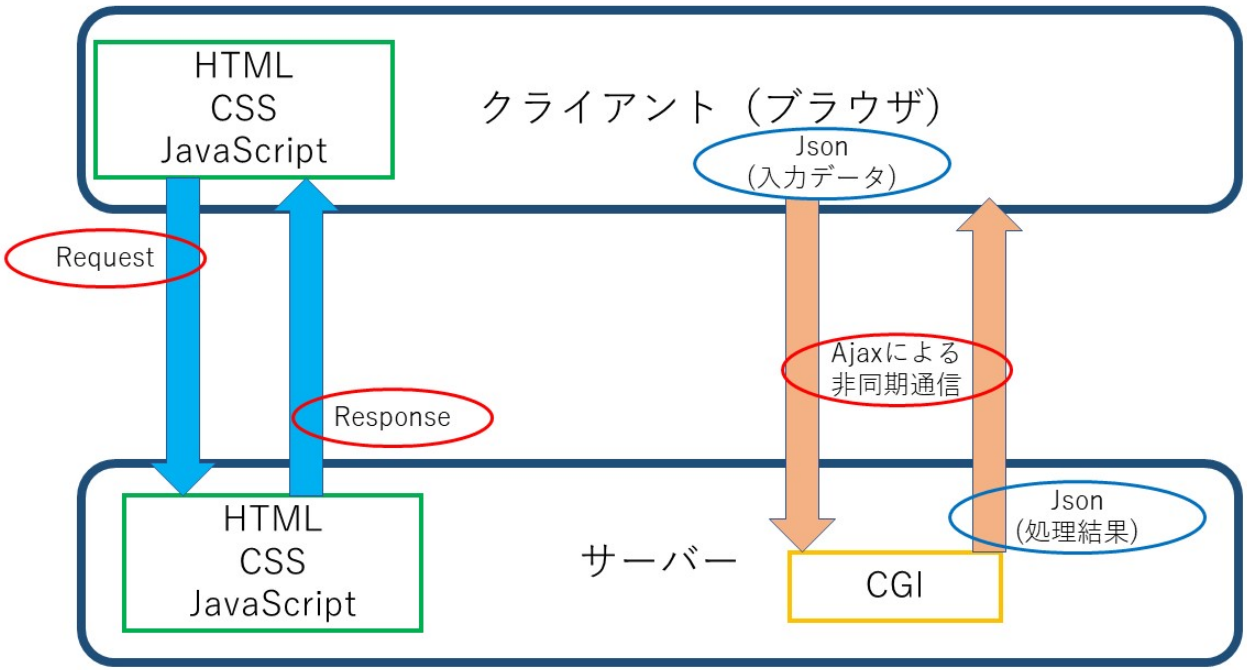

Figure 8: 処理の遷移

4.3 機械学習ライブラリの導入

Common Gateway Interface(CGI)

Ajax による通信から先のバックエンドでは，CGI と呼ばれるプログラムを使って，クライアントから送られたデータを処理して、結果をクライアントに返す.
CGI は python で記述する．python で記述することによって，数理処理用ライブラリやスクレイピング用ライブラリ，マルチメディア用ライブラリなど多岐に渡るライブラリを使用することができる.
本研究では，scikit-learn と呼ばれる機械学習ライブラリを使用した.

Figure 9: python ライブラリの一例

4.4 CSV ファイルの読み込みブロックの作成

入力データの形式

入力データは，データの入力形式をあらかじめ統一し，拡張子は CSV とすることにした．入力形式としては，以下のようにした．

- セルの 1 行 1 列目に入力データ数
- 2 行 1 列目に説明変数，2 行 2 列目に目的変数
- 3 行目にラベル名
- 4 行目に基本データ型
- 5 行目以降に 3,4 行目に対応する入力データ

	A	B	C	D	E	F	G	H
1	1599							
2	11	1						
3	a	b	c	d	e	f	g	h
4	float	float	float	float	float	int	int	float
5	7.4	0.7	0	1.9	0.076	11	34	0.9978
6	7.8	0.88	0	2.6	0.098	25	67	0.9968

Figure 10: 入力データの形式例

5.1 作成したブロック

ファイルを読み込むブロック

ファイルを読み込むブロックの見た目は以下の図である．**ファイルを選択してください**の枠内をマウスでクリックすることで，ポップアップでエキスプローラーが開き，そこから入力データとなる CSV ファイルを選択すると，入力データを JSON に置き換えてこのブロック自体に内容を保持させることができる．

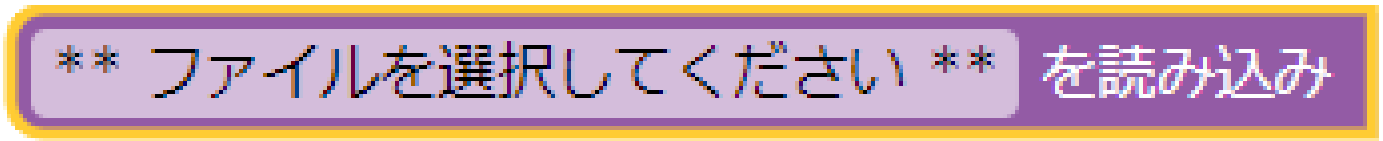

Figure 11: ファイルを読み込むブロック

5.1 作成したブロック

出力結果のダウンロードができるブロック

出力結果のダウンロードを CSV 形式と PNG 形式でダウンロードできるブロックを以下の図のように作成した．このブロックの使い方としては，ブロックの空いたところが入力として，プログラムの実行時に入力部分の処理結果を CSV に変換し，そのファイルのダウンロードが開始されるものとなっている．

Figure 12: 出力結果のダウンロードができるブロック

5.1 作成したブロック

データ分析のブロック

データ分析は線形回帰の計算をするブロックと非線形回帰をするブロック以下の図のように作成した．このブロックの使い方としては，ブロックの空いたところに入力データとなる JSON が含まれるブロックを入れることで実行時に計算の処理が行なわれ，もとの入力データと予測値の結果や回帰係数，切片，決定係数が返ってくるものとなっている．

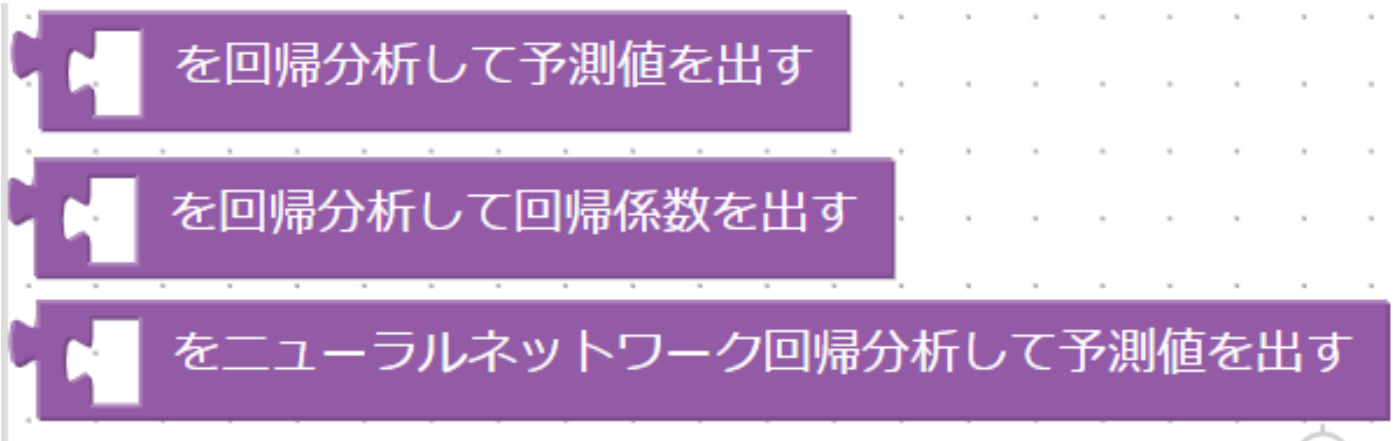

Figure 13: データ分析のブロック

5.1 作成したブロック

クラスター分析のブロック

次に，クラスター分析の計算をしてブロック以下の図のように作成した．このブロックの使い方としては，ブロックの空いたところに入力データとなる JSON が含まれるブロックを入れることで実行時に計算の処理が行なわれ，画像データを base64 型の値にエンコードした値が返ってくるものとなっている．

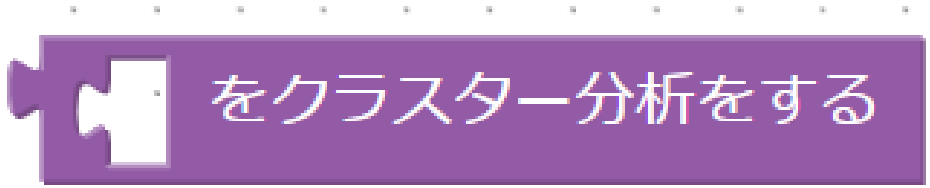

Figure 14: クラスター分析のブロック

5.2 動画

wiki のアップロード制限のため動画は DRIVE を参照してください

考察

ビジュアルプログラミング言語の Blockly を使うことで，プログラミング初心者にも使いやすいものができた．
今回作成したブロックを以下の図のように組み合わせることで，簡単に Blockly で入力データに対して線形回帰の計算や非線形回帰, クラスター分析の計算をし，結果を画像や CSV ファイルでダウンロードできるようになった．

Figure 15: 作成したブロックを組み合わせた図

6. おわりに

まとめ

本研究ではビジュアルプログラミング言語によるデータ解析システムの開発をした.Ajax 通信を使って Blockly 上で線形回帰，非線形回帰，クラスター分析の計算ができるようになった.

今後の課題

- データを分析する手法を調べ，そのブロックを作成していく.
- 今回作った線形回帰のブロックを編集してオプションで単回帰分析か重回帰分析をブロックで選択できるようにする.
- HTML 上に結果をグラフ表示できるようにしたい
- 外部に公開できるようにしたい.