

令和3年3月12日
物価統計室

ウェブスクレイピング及び機械学習を活用した衣料品ネット価格収集の検討状況

1 背景

消費者物価指数 2020 年基準改定においては、POS データ及びウェブスクレイピング技術を活用し、教養娯楽用耐久財及び旅行サービスについて、新たにネット販売価格を把握することとしている。

ネット購入の割合はこのところ年々拡大する傾向にあり、足下では新型コロナウイルス感染症の拡大を契機に急伸している状況にある¹。POS データやウェブスクレイピングの活用は、調査員を介さない非接触型の価格収集方式という意味においても有用であり、今後も取組を拡大していくことが必要である。

衣料品は、ネット販売の市場規模及び購入割合が比較的大きく、これまで試験的にウェブスクレイピングによる価格収集を行い、ネット価格採用に当たっての課題検討を行ってきた（平成 31 年 3 月 6 日 第 13 回物価指数研究会）。衣料品のウェブスクレイピングデータには、求める衣料品のほかに大量の関連商品が混在するため、この中から特定商品あるいは同等商品の抽出が必要となるが、商品コードや名称等が存在しないことが多いためにこれらをキーとした機械的なフィルタリングは難しく、商品説明文や画像から人手で抽出することもおよそ現実的ではない。

表 1 近年のネット販売の進展状況（物販系分野）

	2017 年		2018 年		2019 年	
	EC 市場規模 (億円)	EC 化率 (%)	EC 市場規模 (億円)	EC 化率 (%)	EC 市場規模 (億円)	EC 化率 (%)
衣類・服装雑貨等	16,454	11.5%	17,728	13.0%	19,100	13.9%
生活家電、AV 機器、 PC・周辺機器等	15,332	30.2%	16,467	32.3%	18,239	32.8%
食品、飲料、酒類	15,579	2.4%	16,919	2.6%	18,233	2.9%
生活雑貨、家具、 インテリア	14,817	20.4%	16,083	22.5%	17,428	23.3%
書籍、映像・音楽ソフト	11,136	26.4%	12,070	30.8%	13,015	34.2%
化粧品、医薬品	5,670	5.3%	6,136	5.8%	6,611	6.0%
自動車、自動二輪車、パー ーツ等	2,192	3.0%	2,348	2.8%	2,396	2.9%
事務用品、文房具	2,048	37.4%	2,203	40.8%	2,264	41.8%

出典：経済産業省「電子商取引に関する市場調査」

¹ 家計消費状況調査結果（総務省）によると、緊急事態宣言が発出された 2020 年 4 月以降、ネット購入の利用世帯の割合は大きく増加し、同年 5 月には調査開始以降初めて 5 割を超えている。

このため今回、今後のウェブスクレイピング拡大に当たっての課題検討に資するため、衣料品のウェブスクレイピングデータについて、商品説明文から自動的に同等商品にカテゴリ分類する機械学習モデルを構築し、その精度の検証や実運用の可能性を考察した。

2 先行研究

衣料品のウェブスクレイピングデータを機械学習によりカテゴリ分類する試みは、オランダ、英国等において行われている。

オランダ²は、単独のオンラインサイトから得た衣料品データ（COICOP 分類レベルのカテゴリが既に付与された約 32 万件）について、短い商品説明文（short description）などから機械学習によりカテゴリを予測し、どの程度一致するかを試算している。複数の機械学習アルゴリズムを比較し、商品分類に最適なモデルはロジスティック回帰（精度：92%）としている。

英国³は、小売業サイトから得た衣料品データ（約 5 万件）について人手でカテゴリを付し、4～8 語からなる商品名から機械学習によりカテゴリ（「男性用シャツ」「婦人用コート」など）を予測する研究を行っている。複数のアルゴリズムを比較し、勾配ブースティング（精度：79%）の手法が最も精度が高いとしている。

以下で示す試みも、オランダや英国の研究を参考としている。大きな違いは、①日本語は単語で区切られていない言語であるため、諸外国の研究に比べて難易度がやや高いと考えられること、②以下の分析では、オランダや英国に比べ長い商品説明（単語数にして 100～400 程度）を元にかなり細かいカテゴリ分類を試みている点が挙げられる。

3 マスターデータの取集

一連の分析の元となるウェブスクレイピングデータは、衣料品を販売する複数のウェブサイト（7 サイト）から取集した、「子供用ズボン」（約 30 万件）及び「運動靴」（約 50 万件）の商品情報である。データには、価格情報に加えて商品の特徴を示す説明文が含まれている。取集に当たっては、小売物価統計調査の基本銘柄⁴と可能な限り同品質の商品を取集できるよう条件を設定した。

取集価格数は表 2 のとおりである。なお本データは、第 13 回物価指数研究会（平成 31 年 3 月 6 日）において示したデータ（2018 年に取集）の一部である。

² “A comprehensive view of machine learning techniques for CPI production”（2019）において、ロジスティック回帰、ナイーブベイズ、ランダムフォレスト、サポートベクターマシン、勾配ブースティング決定木、ディープラーニング（RNN、アテンション層、CNN）が比較されている。

³ “Automated classification of web-scraped clothing data in consumer price statistics”（2020）において、ロジスティック回帰、サポートベクターマシン、勾配ブースティング決定木、決定木、ランダムフォレストが比較されている。

⁴ 男児用ズボン：「長ズボン，〔素材〕「綿 100%」又は「綿 95%以上・化学繊維混用」，〔サイズ〕140 又は 150，普通品」、運動靴：「大人用，スニーカー，〔甲〕合成繊維・合成皮革，〔タイプ〕ひも又はマジックテープ，〔サイズ〕23.0～26.0cm，中級品，「スポルディング」又は「マックスランライト（ダンロップ）」

表2 マスターデータの件数

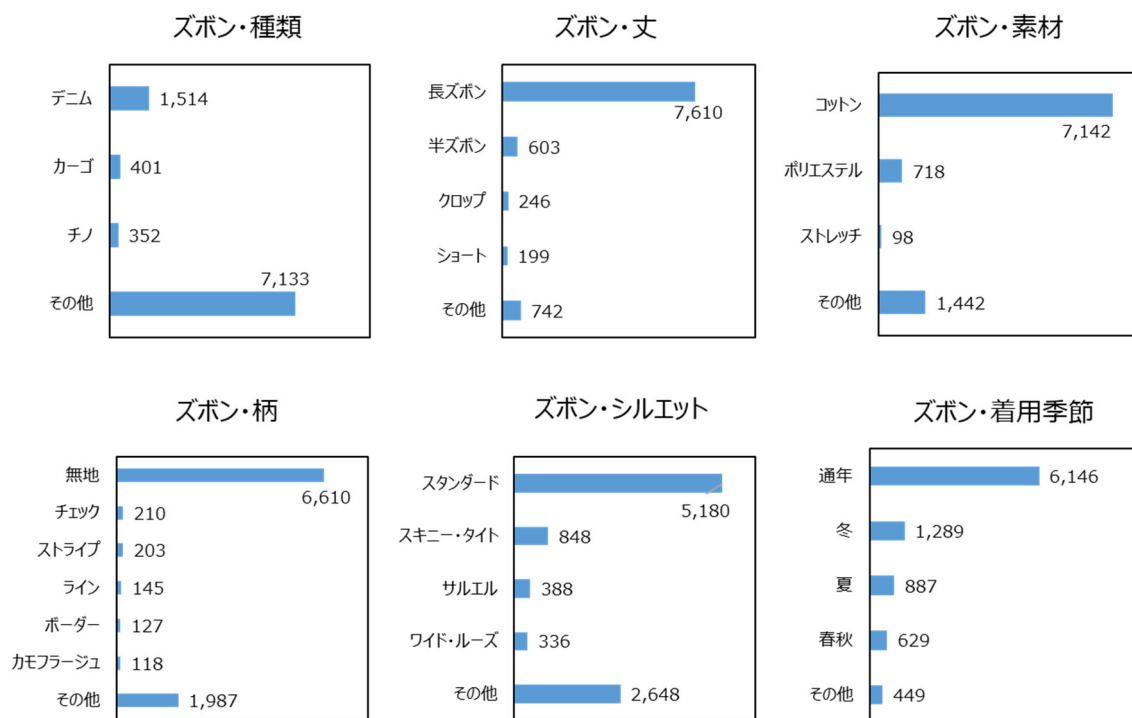
品目名	収集時期	2018年 価格数(概数)							計
		6月	7月	8月	9月	10月	11月	12月	
子供用ズボン	1～12月	27,200	55,300	60,900	63,200	50,900	45,900	44,900	348,300
運動靴	1～12月	28,200	91,700	100,600	95,600	75,000	65,400	70,500	527,000

4 教師データの作成

商品説明文から商品の特徴を推定する機械学習モデルは、教師あり学習により構築する。その教師データ（学習用・検証用データ）を作成するため、マスターデータの一部について、人手で商品説明文から当該商品のカテゴリを判断し、そのラベルを付与する作業を行った。

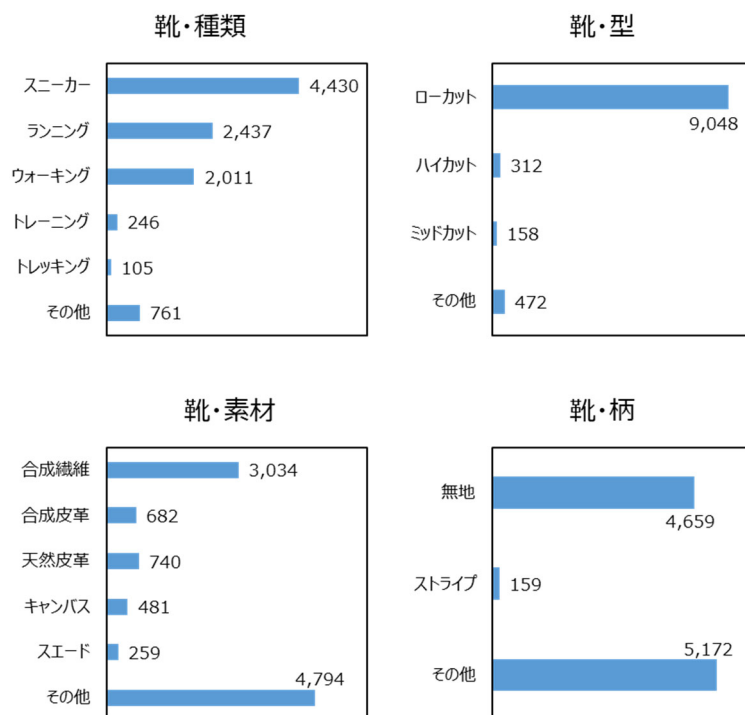
ラベル付け作業は民間事業者への委託研究⁵の一環として、クラウドソーシングの方法により「子供用ズボン」及び「運動靴」（各1万件程度）について行った。ラベルは、小サンプルによる事前検討を経て、ズボン・靴の種類や素材などについて、商品の特徴づけるものを適宜設定した。ラベル付けの結果は図1のとおりである⁶。

図1 ラベル付けの結果



⁵ 「人工知能を活用した衣料品ネット通販価格のデータ分析業務」（2019年10月～2020年3月）

⁶ ラベル付けは各商品レコードに対しクラウドワーカー3名以上によって行っており、判定が分かれた場合には多数決によって当該商品のラベルを確定した。なお委託研究では、他者との回答の一致率が低く信頼度が劣ると判断されるクラウドワーカーの回答は除外するなどの処理も行っている。



5 機械学習

4の教師データを用いた教師あり学習により、商品説明文からカテゴリを予測するモデルを以下の方法で構築した。

(1) 商品説明文のクリーニング

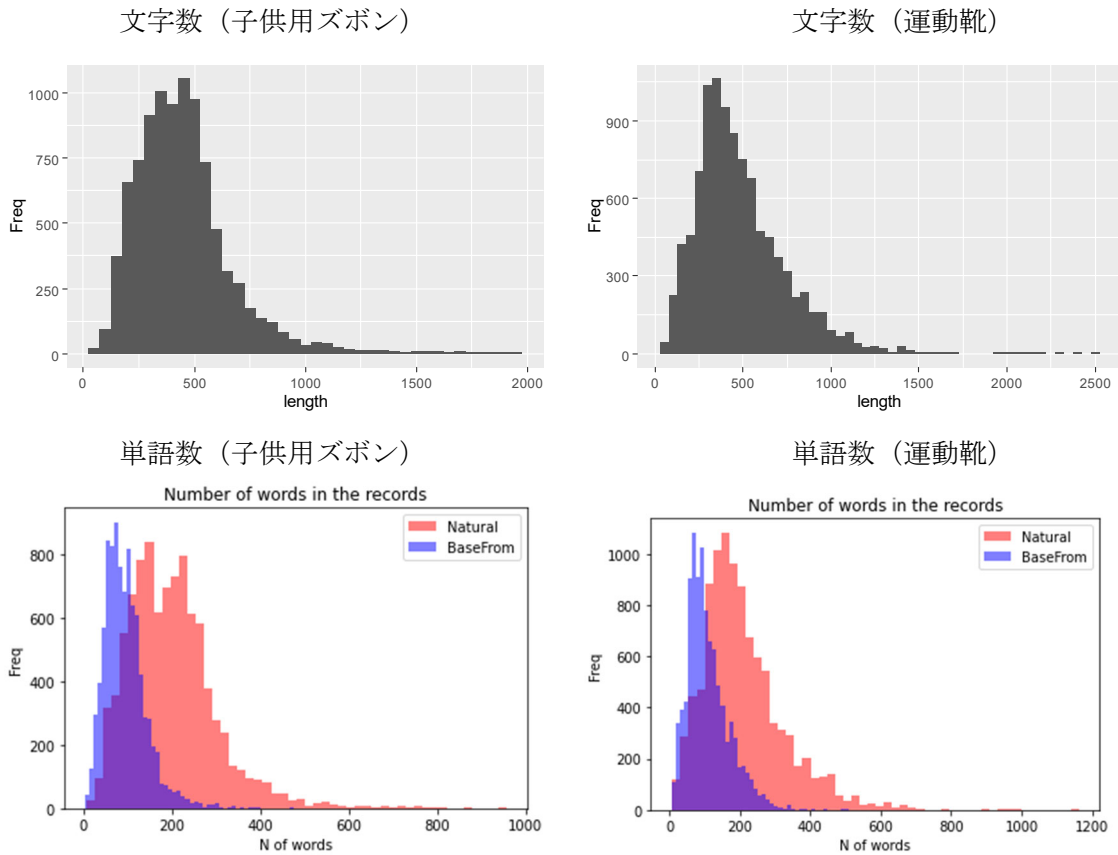
商品の説明文の長さはばらつきがあるが、200字～500字程度、単語数で100～400語程度の商品が多くなっている。説明文には、商品の特徴の説明の他、決済方法や発送方法などの情報、各種広告、記号などが多く含まれている。

機械学習に当たっては、モデルのシンプル化による精度・解釈性の向上や学習時間短縮のため、不要な記号の削除に加え、特定の品詞（名詞、動詞、形容詞及び副詞）のみを抽出しその原型に変換することによって、特徴量数（機械学習に用いる単語数）を圧縮した。クリーニングの結果、特徴量数は、子供用ズボンは18,569から3,496に、運動靴は20,599から4,942となった。

なお単語の分割には、汎用の形態素解析エンジン MeCab⁷を用いた。

⁷ 委託研究では、sentencepiece（日本語の文法とは無関係に学習用データから効率的に文章を分割するアルゴリズム）が用いられたが、本分析では精度向上が見られなかったことなどから、採用していない。

図2 説明文の文字数・単語数の分布



注：Natural は原文、BaseForm はデータクリーニング後（品詞限定・原形化含む）のもの

表3 商品説明文に使用されている単語及び頻度（上位のもの）

子供用ズボン				運動靴			
原文		クリーニング後		原文		クリーニング後	
：	63,038	する	18,452	の	76,121	する	37,546
の	48,937	パンツ	18,398	。	56,514	素材	13,457
。	43,569	丈	16,273	、	55,495	スニーカー	12,675
、	40,485	綿	13,161	：	53,376	シューズ	11,650
（	32,194	素材	12,666	を	48,976	いる	10,930
に	30,464	ウエスト	12,647	に	45,622	れる	9,675
）	29,524	ある	9,637	は	29,246	配送	9,386
/	26,052	配送	9,305	（	27,955	サイズ	9,355
が	26,002	即日	9,007	で	26,549	即日	9,034
は	25,444	商品	8,140	た	26,364	履く	8,292
ます	21,285	サイズ	7,982	が	26,233	メンズ	8,007
で	19,592	ポリウレタン	7,874	し	23,426	アッパー	7,507
を	19,500	場合	7,252	て	23,329	送料	7,147
て	18,998	いる	6,737	/	21,710	ある	6,873
パンツ	18,398	送料	6,538	）	21,509	商品	6,536
総	16,362	ゴム	6,366	な	21,372	底	6,010
丈	16,273	股下	6,010	ます	20,602	タイプ	5,873
.	15,527	タイムセール	5,895	と	17,724	くださる	5,562
】	14,806	ポケット	5,884	・	17,146	デザイン	5,488
【	14,358	くださる	5,449	も	13,620	場合	5,447

(2) 特徴量ベクトルの算出

クリーニングした説明文（単語の集まり）からカテゴリを推測する機械学習モデル構築のため、これを数値ベクトルに変換する必要がある。

変換に当たっては、文章中の単語の存否によってベクトルを作成する最もシンプルな BoW (Bag of Words)、それらの単語の頻出頻度に応じて重みをつける TF-IDF (Term frequency - Inverse document frequency)、単語の意味を勘案して数百次元程度のベクトルに埋め込む方法 (Word2Vec、Fasttext など) があるが、ここではシンプルさと精度を考慮し、TF-IDF の方法を用いた⁸。

(3) 機械学習の方法

教師あり機械学習のアルゴリズムには多くの選択肢があるが、今回は先行研究でも用いられている、ロジスティック回帰（オランダの先行研究で最適とされている）、カーネルサポートベクターマシン、ランダムフォレスト、勾配ブースティング決定木（英国の先行研究で精度が高いとされている）を比較検討することとした⁹。

学習は、教師データを 8 : 2 の比率で学習用データとテスト用データに分割し、学習用データに対し 5 回の交差検証法によって行った。得られたモデルについてテスト用データにおける正解率¹⁰を算出し、それを各アルゴリズムで比較した。各アルゴリズムとも、シンプルな二値分類モデル（最多のラベルとそれ以外への分類）と多値分類モデルを試算した。

これらには、モデルのパフォーマンスを左右する各種パラメータがあるが、その調整を厳密に行うには相当の時間を要する。今回は、将来的な利用可能性を模索するための予備的な検討であり、分類精度を究極に求めるものではないことから、パラメータ調整は簡易的に行うのみとした。

なお、これらのアルゴリズムの実装には、汎用言語 Python 及びオープンソースの機械学習ライブラリ scikit-learn を主として用いた。

(4) 機械学習の結果

各アルゴリズムによる正解率及び演算時間¹¹は、表 4 及び 5 のとおりである。正解率は、アルゴリズムによって若干な差が見られるが、おおむね良好な成績（一部のカテゴリ¹²を

⁸ 委託研究では単語埋め込み (Word2Vec) 及び文章ベクトル変換 (SCDV) が行われたが、今回分析では目立ったこれらの有効性は見出せなかった。

⁹ 委託研究においては、k 近傍法（正解率平均 68.1%）、カーネルサポートベクターマシン（同 67.6%）、勾配ブースティング（同 78.8%）について比較検証された。この他、ニューラルネットワークを用いたより複雑なモデルも考えられる。念のため簡易的に、単語埋め込み及びニューラルネットワークによる複数方法 (RNN、LSTM、CNN) も試算したが、大きな精度向上は得られなかった。

¹⁰ 一般にモデルの評価指標としては、正解率 (accuracy) のほか、適合率 (precision)、再現率 (recall)、F 値があるが、ここでは簡易に正解率のみを比較した（他の指標で比較してもあまり結論は変わらない）。

¹¹ Windows 10 Pro 64bit、Intel(R) Core(TM) i3-3220 CPU @ 3.30GHz、RAM 8.00GB の環境にて実施。

¹² 「ズボン・シルエット」や「靴・柄」は他に比べ正解率が低い。例えば「靴・柄」については、説明文

除いて8～9割程度の正解率)が得られた。一方で演算時間は、ロジスティック回帰が高速(10秒～1分程度)であるが、カーネルサポートベクターマシンはかなり時間を要している(10～40分程度)。ただし演算時間にはパラメータ調整も含まれており、ロジスティック回帰以外については簡易な方法¹³をとったため、この厳密性を高めると更に多くの時間を要することに留意が必要である。

ロジスティック回帰は、係数の大小によって簡易的にカテゴリ判定に影響した単語を見出すことができ、アルゴリズムがおおむね期待どおりに働いている様子がうかがえる(表6)。他のモデルについては(判定に影響した単語を出力することが可能なものもあるが)アルゴリズムが若干複雑であり、ややブラックボックスとなっている。

表4 機械学習アルゴリズム(二値分類)の比較

		子供用ズボン						運動靴			
		種類	丈	素材	柄	シルエット	着用季節	種類	型	素材	柄
		デニム =1	長ズボン =1	コットン =1	無地 =1	スタンダード =1	通年 =1	スニーカー =1	ローカット =1	合成繊維 =1	無地 =1
正解率	ロジスティック回帰	0.931	0.896	0.829	0.836	0.727	0.884	0.789	0.954	0.915	0.565
	カーネルSVC	0.942	0.880	0.823	0.823	0.701	0.864	0.763	0.950	0.902	0.541
	ランダムフォレスト	0.906	0.885	0.807	0.789	0.697	0.848	0.787	0.933	0.895	0.527
	勾配ブースティング	0.943	0.896	0.849	0.838	0.718	0.893	0.794	0.956	0.919	0.535
	ロジスティック回帰	10s	10s	13s	12s	12s	12s	12s	11s	14s	12s
演算時間	カーネルSVC	14m7s	15m32s	19m44s	19m52s	24m55s	19m28s	30m28s	15m55s	22m23s	37m12s
	ランダムフォレスト	47s	51s	48s	51s	59s	55s	1m18s	45s	1m21s	58s
	勾配ブースティング	22s	22s	22s	22s	23s	23s	28s	29s	27s	27s

表5 機械学習アルゴリズム(多値分類)の比較

		子供用ズボン						運動靴			
		種類	丈	素材	柄	シルエット	着用季節	種類	型	素材	柄
正解率	ロジスティック回帰	0.918	0.895	0.810	0.819	0.683	0.865	0.767	0.943	0.852	0.526
	カーネルSVC	0.927	0.876	0.792	0.812	0.670	0.843	0.723	0.946	0.827	0.530
	ランダムフォレスト	0.850	0.872	0.784	0.758	0.611	0.803	0.733	0.927	0.801	0.510
	勾配ブースティング	0.933	0.881	0.802	0.818	0.671	0.858	0.744	0.940	0.845	0.524
	ロジスティック回帰	26s	30s	31s	40s	36s	34s	45s	38s	60s	1m41s
演算時間	カーネルSVC	25m9s	28m36s	27m53s	39m30s	34m28s	30m34s	45m14s	30m39s	40m10s	41m3s
	ランダムフォレスト	47s	55s	51s	56s	56s	1m8s	1m28s	46s	1m44s	59s
	勾配ブースティング	1m8s	1m15s	1m13s	1m52s	1m25s	1m37s	2m2s	1m21s	2m12s	1m2s

に柄の言及がない商品についてクラウドワーカーによって「無地」又は「その他」と判定が割れており、ラベルの定義やラベル付ルールの明確化が重要なことが示唆される。

¹³ ロジスティック回帰については正則化(L1, L2)及びパラメータCについてグリッドサーチにより調整。カーネルサポートベクターマシン(SVC)及びランダムフォレストはパラメータの分布を仮定した簡易な調整。勾配ブースティングのパラメータはデフォルト(フレームワークLightGBMを使用)。

表6 カテゴリ判定に影響した単語の例（二値分類モデル）

	子供用ズボン						運動靴			
	種類	丈	素材	柄	シルエット	季節	種類	型	素材	柄
	デニム=1	長ズボン=1	綿=1	無地=1	スタンダード=1	通年=1	スニーカー=1	ローカット=1	合成繊維=1	無地=1
正の係数 上位	デニム, ニット デニムパンツ, ダメージ, mp, denim, 子供 服, サックス, デニムワイドパ ンツ, デニムロ ングパンツ, ジーンズ	ロング, 長ズボ ン, 秋冬, レギ ンス, スキニー パンツ, 先, ウオッシュ, ポ リウレタン, シヤ ギー, 起毛	綿, コットン, 柔らかい, 春 夏, 夏, 裾リブ, 加工, 革, bl, き口	シンプル, 保 温性, 合皮, 付属, xl, ア ジャスターゴ ム, キャメル, 活躍, ng, シ ルエット	スウェットパン ツ, カーゴパン ツ, デニム, チ ノパン, チノパ ンツ, クライミ ング, カット, ク ロップドパンツ, 開く, カーゴ	表記, レイヤ ード, 夏秋, 伸縮 性, サロベツ, 使える, チノ, ストレッチデ ニムパンツ, られ る, クロップド パンツ	スニーカー, メ ンズスニー カー, カースポ ル, デイニング, カット, 滑る, レザー, white, 路面, 表記, ス リッポン, mens	ランニング, ウォーキング, シューズ, ナイ ロン, ローカッ トモデル, オッ クスフォード, 合成繊維, so, ゆったり, lo	合成繊維, ナ イロン, ポリエ ステル, ポリエ ステルソール, クラシックコ ル, テツナイロ ン, ultk, コーデ デュラナイロ ン, ゴム, クッ ションニ ング, 相当	通学, 防水, ホワイト, やす い, 在庫, タイ ムセール, 不 可, インソ ール, black, spalding
負の係数 上位	カーゴパンツ, スウェットパ ンツ, グレー, カーキ, チノ, コーデュロイ, ガウチョパン ツ, ベージュ, dm, 本物	ハーフパンツ, ショートパン ツ, クロップド パンツ, 七分, ハ ーフ, クロップ ド, 夏, 分, ショ ーツ, ダンブ	ポリエステル, ナイロン, 麻, ストレッチ, ス ウェット, シヤ ギー, コー デュロイ, レー ヨン, メッシュ, 裏	ストライプ, チェック柄, チェック, 迷 彩, 柄, ボー ダー, ライン, 総柄, ボー ダー柄, サイ ドライン	スキニーパ ンツ, サルエル パンツ, ワイド, スキニー, テーパードシ ルエット, ルー ズ, サルエル, 細身, スキ ニーシルエット, バルーン	秋冬, 春夏, ショートパン ツ, ハーフパ ンツ, ワイド, 秋, コーデュ ロイ, 冬, 春秋, 夏, 春	ランニング, ウォーキング, トレニング, ジョギング, ラ ンナー, シューズ, モダ ンランニング シルエット, ト レニングシュ ーズ, 極限, ダ ンロップ	ハイカット スニーカー, ミ ッドカット, hi, ハイ, mid, カッ ト, airjordan, ミッドカッ トタイプ, high, シリーズ	合成皮革, 皮 革, 人工, フォ ルム, 天然, ロ ゴマーク, キ ャノンバス, ア クセント, レ ザー, 紳士	柄, スリー, ス トラップス, eq t, staple, 快適 一足, シグネ チャー, あしら う, サイド

6 価格動向の試算

5で構築したモデルが同等商品の抽出にどの程度有効かを分析するため、モデルの予測によって特定カテゴリに該当した商品群の価格動向を観察する。

そのため、説明文の表記が比較的整っている大手サイト（1サイト）のデータ（約15万件）に対し、5で構築したロジスティック回帰モデル（多値分類、二値分類）を適用して各商品のカテゴリを予測した。予測結果の分布（多値分類、図3）は教師データ（図1）と似通っており、モデルはおおむね期待どおりに機能していると推測される。なお予測は各々所要60秒程度であった¹⁴。

価格動向を試算する商品群カテゴリとしては、学習精度、予測結果の分布、予想される価格への影響度などを勘案し、子供用ズボンについては「長ズボン」かつ「通年」、運動靴については「スニーカー」を用いることとした¹⁵。より安定的な結果を得るために、カテゴリ予測はより精度の高い二値分類モデルを適用した。ネット価格における外れ値やタイムセールなどの攪乱要因を一定程度コントロールした上で、該当商品群の価格動向（幾何平均）を試算すると、図4のとおりとなった。

¹⁴ 委託研究により得られた勾配ブースティングによるモデル（sentencepieceによる単語分割のための事前学習、Word2Vec及びSCDVの適用を含む）では、全データのカテゴリ予測に5時間以上を要し、実用性の面で大きな課題がある。

¹⁵ 「柄」は価格を左右する品質と考えにくいいため、「素材」は部位によって素材が異なる商品のラベル付ルールが不明確と考えられるため、除外。「ズボン・シルエット」は学習精度が低いため除外。「ズボン・種類」は最多の「デニム」でもレコード数が少なく除外。「靴・型」は安定的な分析結果が得られず除外。

図3 カテゴリ推測結果（多値分類）

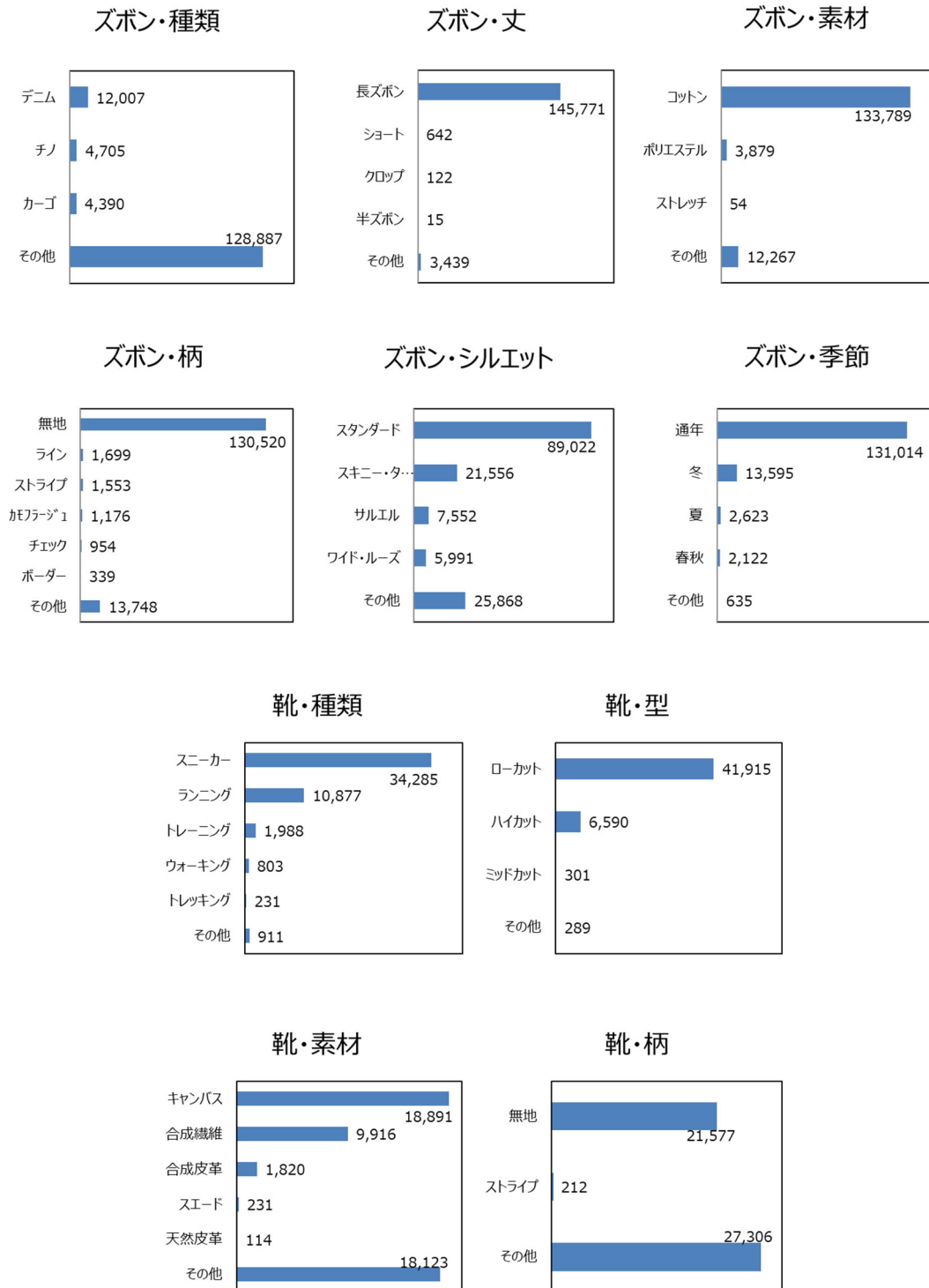
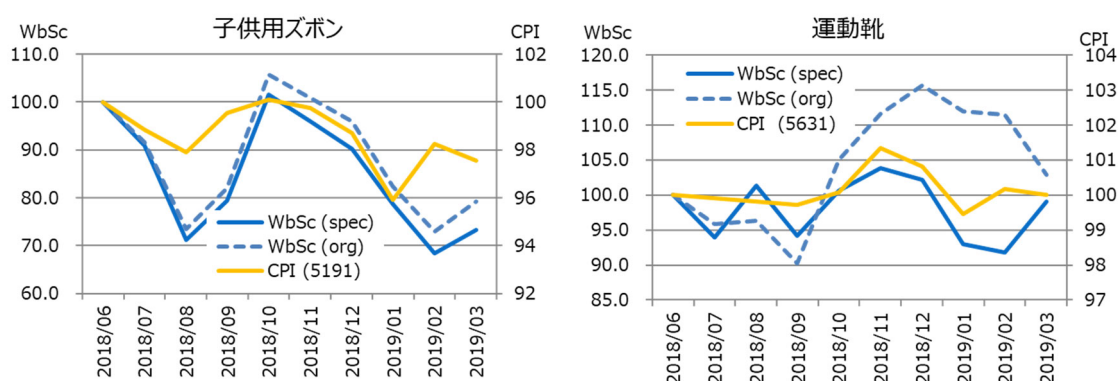


図4 カテゴリ予測の有効性



凡例 WbSc(org) : ウェブスクレイピングデータ (外れ値・タイムセール除去後)

※レコード数 子供用ズボン 58,690、運動靴 9,410

WbSc(spec) : うちカテゴリ該当商品群 (子供用ズボン:「長ズボン」「通年」、運動靴:「スニーカー」)

※レコード数 子供用ズボン 48,197、運動靴 3,624

子供用ズボンについては、「長ズボン」かつ「通年」に限定しても、価格動向に目立った変化は見られなかった。他方、運動靴については「スニーカー」に限定することで、価格動向が CPI の動きに接近した。この分析はかなり簡易なものであり、早計な判断は難しいが、運動靴については同等商品の抽出に有効であった可能性がある。

ただし、子供用ズボンの場合も運動靴の場合も、変動幅については CPI と大きな差が見られる。試算値の変動幅が大きいのは、単独サイトの商品を試算対象としたことがその要因の一つと考えられ、更なる研究の改善が必要である。

7 まとめと今後の課題

(1) 機械学習

今回行った機械学習によるカテゴリ分類は、一部を除いて8～9割程度の正解率であり、一定程度確からしいモデルが得られた。中でも、ロジスティック回帰による方法は、演算時間が短くメンテナンスが容易で、説明もしやすいことから、今後の有効性・実用性が期待できると考えられる (オランダにおける先行研究とも符号する)。引き続き、他の品目への適用可能性や分析手法の改良など、検討を続けることとしたい。

また、機械学習アルゴリズムを適用する前提として、教師データが必要であり、その質と作成コストが大きな課題と考えられる。今回分析に用いた教師データは、民間事業者の管理下で行われたクラウドソーシングによるものであるが、同等商品を区別するためのカテゴリの検討、ラベル付ルールの明確化、ラベラーのコントロールが課題である。他方、英国の先行研究では、教師データの作成は ONS 職員が行っている。当室には、小売物価統計調査の調査品目 (銘柄) の維持・管理を担当する専門職員が在籍しており、この専門知識を活用することも考えられる。引き続き体制面も含めて、

機械学習の活用可能性の研究を進めていくこととしたい。

(2) ネット価格の価格動向の算出方法

今回、機械学習により得られたカテゴリを用いて、同等商品群の価格動向を簡易に試算したところ、一定の有効性は示唆されたものの、実運用にはまだ相当の研究分析を要することが確認された。

検討すべき課題としては、機械学習の手法以前に、ウェブスクレイピングに起因する課題も多いと考えられる。例えばネット価格と店頭価格の違い、対象品目・銘柄の違い、ウェブスクレイピングしたサイトの代表性、外れ値の処理方法（買い手がつかないような価格の判定方法）、セールスの扱いなどについて、複合的な検証が必要と考えられる。

衣料品（衣類・服飾雑貨等）のネット購入割合は、直近では1割強（表1）であり、これまでのところ現行の店頭販売を中心とした価格収集・CPI作成に支障は生じていないが、感染症に伴うネット販売の急増も踏まえ、（今回の分析手法に限らず）ネット価格の反映手法の研究を、引き続き幅広く進めていくこととしたい。