

Application News

No. M276

ガスクロマトグラフ質量分析計

機械学習的解析手法を利用した食品の 品質判別

食品の多くはきわめて多くの成分で構成されており、同じ食品でもその品質は完全に同一ではないことがあります。品質の違いは食品を構成する成分の微妙な違いに由来すると考えられるため、昨今では包括的な品質評価を目的とした、構成成分の網羅分析が注目されています。一方で、味、におい、劣化の度合いなど、人による主観的な食品の性質を構成成分から予測・判別するためには、構成成分と主観的な性質の関係について、予めその対応が明らかな既知のサンプル群から学習を行い、その学習結果を元に未知サンプルの予測・判別を行うという手法が有効と考えられます。

本アプリケーションニュースでは、牛肉の品質に関して、適切に冷蔵保存されているものと、40℃環境に3時間暴露され劣化が進んだと考えられるものを準備し、それらを200℃に加熱したときの揮発成分の分析結果から判別が可能かどうか検証しました。事前にそれぞれのサンプルの既知データを識別器に学習させて、これを品質の定義とし、その後その結果を元に未知データを判別させ、その正答率を算出しました。クロマトグラムの比較や、ピーク面積値の主成分分析だけでは判別が難しいサンプルでも、識別器としてサポートベクターマシン (SVM) を使用した場合、95.8%の確率で正答することができました。

T. Sakai

■ サンプルの準備

牛肉のさまざまな部位の赤身に関して、適切に冷蔵保存されているもの (4℃サンプル) と、40℃環境に3時間暴露され劣化が進んだと考えられるもの (40℃サンプル) を準備しました。サンプルの概観を図1に示します。

これらの赤身に関して、20±3 mg を測定用バイアルに取ったものを、116本 (4℃サンプル 58本、40℃サンプルを58本) 準備し、分析に供しました。



図1 左：適切に冷蔵保存されているもの (4℃サンプル)
右：40℃環境に3時間暴露されたもの (40℃サンプル)



図2 AOC-6000の概観

■ Solid Phase Microextraction (SPME) による分析

バイアルを200℃で15分加熱し、発生する気体をSPMEで捕集してスキャン分析しました。サンプル数が多いので、注入には自動でSPMEでの捕集・脱着が可能なAOC-6000を使用しました。分析条件を表1に示します。

トータルイオンクロマトグラムの比較からは、各サンプルを特徴付けるピークを見つけることはできませんでした。

表1 測定条件

SPME fiber	: Divinylbenzene/Carboxen/Polydimethylsiloxane (DVB/CAR/PDMS)
Incubation Temp.	: 200℃
Incubation time	: 15 min
Agitator	: 250 rpm
Desorb time	: 1 min
Column	: SLB™-5MS (30.0 m×0.25 mm, 0.25 μm, Sigma-Aldrich Co. LLC)
Injection mode	: Split
Split ratio	: 5:1
Injection port Temp.	: 280℃
Oven Temp. program	: 60℃ (1 min) → (20℃/min) → 200℃ → (8℃/min) → 320℃ (5 min)
Flow control	: Linear velocity (50.0 cm/sec)
Purge flow rate	: 3 mL/sec
Interface temperature	: 200℃
Ion source temperature	: 250℃
Event time	: 0.3 sec

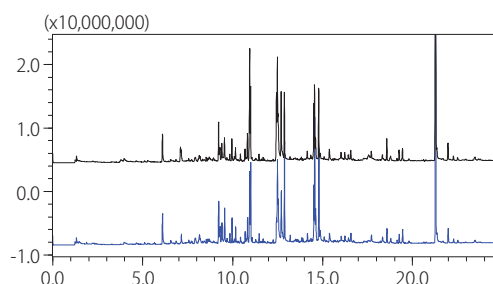
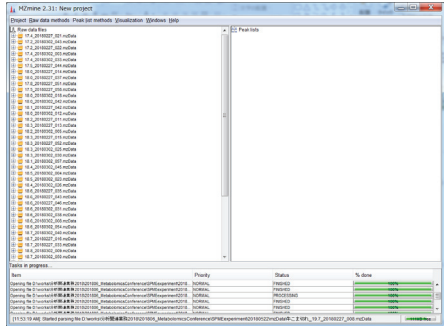


図3 トータルイオンクロマトグラムの例
黒：4℃サンプル 青：40℃サンプル

■ ピークの抽出

質量分析データの解析ソフトウェアである MZmine 2 (Ver. 2.32) を使用して、各 MS クロマトグラムからピークの抽出を行いました。9318 ピークが検出されましたが、全 116 データの中でひとつでも欠損値 (未検出) があるものを削除し、全部で 200 ピークの面積値を説明変数として使用しました。この 200 ピーク×116 データの行列を解析に用いました。



1	freshness	RT1.32	RT1.32021RT3.9811RT4.6484RT6.1108RT6.1111RT6.1113RT6.11
2	17.2_20180227_022	0	142655 20447 38003 10300 1277499 186270 65360 120
3	17.2_20180302_043	1	123769 17718 12369 6486 704254 99777 35242 44
4	17.4_20180227_021	1	164652 24732 24104 10174 1103040 155369 53255 74
5	17.4_20180302_003	1	184356 26071 44966 8371 1636468 225575 84552 103
6	17.4_20180302_033	1	157025 22159 41251 7202 1460463 209730 75402 100
7	17.5_20180227_044	0	150354 20934 7410 9279 469350 66412 23548 30
8	17.5_20180227_058	0	145442 20625 27545 8906 1173658 167106 60198 92
9	17.8_20180227_051	1	137218 20572 31183 9732 1554441 223310 77004 129
10	18.0_20180227_014	0	129346 18402 39322 12398 1499886 217903 79206 171
11	18.0_20180227_037	1	215698 28841 27146 8363 1376918 194752 66472 91
12	18.0_20180302_012	0	121548 17959 42813 7906 1492706 211689 74417 118
13	18.0_20180302_018	0	197382 29289 41028 7643 1297863 177850 63503 82
14	18.0_20180302_042	0	128460 17815 36365 6714 1114779 165519 58623 76
15	18.1_20180227_042	0	17863 25302 29941 8814 1087582 152226 56403 78
16	18.1_20180302_057	1	193556 26762 19093 7126 831226 119570 43728 54
17	18.2_20180227_011	1	181500 25748 54582 13409 1881440 263098 91677 147
18	18.2_20180302_005	1	159436 21548 41616 7603 1199737 168974 60174 79
19	18.2_20180302_013	1	146464 26265 78065 8701 1164335 140183 64666 71

図4 MZmine 2 の操作画面 (上) および、各データにおける各ピーク面積値のデータ行列 (下)

■ 主成分分析

各サンプルに特徴的なピークがないことを確かめるため、このデータ行列を主成分分析に供しました。図5にスコアプロットを示します。主成分分析では、ふたつのサンプル群を区別することは難しいことがわかりました。

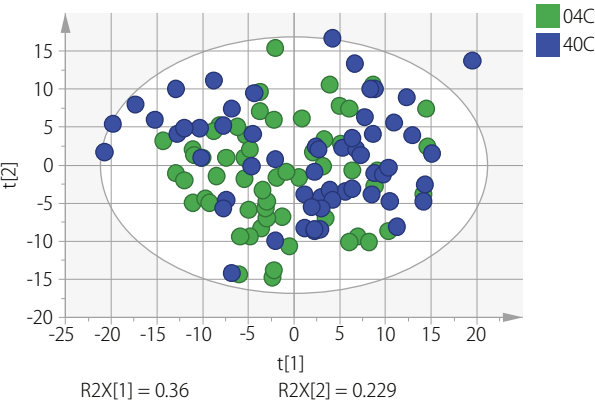


図5 主成分分析のスコアプロット

SLB は、Sigma-Aldrich Co. LLC の商標です。
その他、本資料に掲載されている会社名、製品名、サービスマーク、およびロゴは、各社の商標および登録商標です。
なお、本文中には TM、®マークを明記していない場合があります。

株式会社 島津製作所

分析計測事業部
グローバルアプリケーション開発センター

初版発行：2019 年 1 月
島津コールセンター 0120-131691
(075) 813-1691

※本資料は発行時の情報に基づいて作成されており、予告なく改訂することがあります。
改訂版は下記の会員制 Web Solutions Navigator で閲覧できます。
<https://solutions.shimadzu.co.jp/solnavi/solnavi.htm>

■ 学習データの準備

全 116 サンプルを、92 サンプルで構成されるトレーニングセットと、24 サンプルで構成されるテストセットにランダムに分けました。トレーニングセット、テストセットともに、含まれる 4℃サンプルおよび 40℃サンプルは同数となるようにしました。識別器に対してトレーニングセットで学習を行った上で、テストセットの判別を行いました。

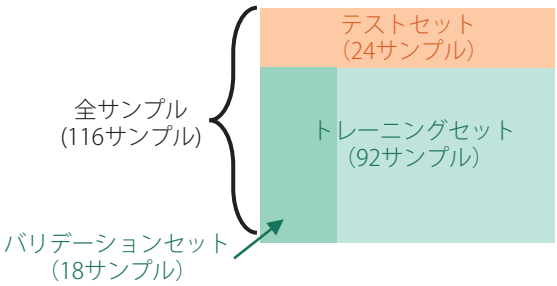


図6 データセット分割の模式図

■ サポートベクターマシンによる判別

識別器としてサポートベクターマシン (SVM) を準備しました。サポートベクターマシンは、Python 3.6 上で scikit-learn (ver.0.19.1) を用いて実装しました。92 サンプルのトレーニングセットをさらにバリデーションセット 18 データとトレーニングセット 74 データに分け、その中で SVM のハイパーパラメータである “C” と “gamma” をクロスバリデーションにより最適化しました。カーネルは “rbf” を採用しました。

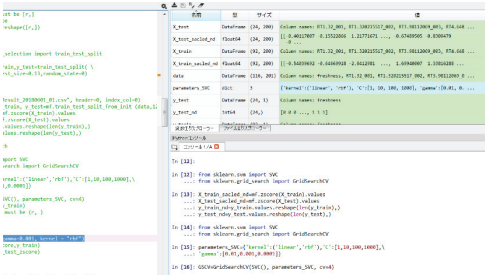


図7 Python 3.6 および scikit-learn による SVM の実装

トレーニングセット内で最適化されたハイパーパラメータを使用し、テストセットである 24 サンプルを判別させました。その結果 24 サンプル中 23 サンプルについて、正しく判別することができました。結果を表2に示します。表2では、4℃サンプルを Positive、40℃サンプルを Negative と表現しています。

表2 テストセット 24 サンプルの判別結果

	True	False
Positive	12	0
Negative	11	1
Precision	95.8 %	

会員制情報サービス「Shim-Solutions Club」にご登録ください。
<https://solutions.shimadzu.co.jp/>
会員制 Web の閲覧だけでなく、いろいろな情報サービスが受けられます。

＞ アンケート

関連製品 一部の製品は新しいモデルにアップデートされている場合があります。



＞ AOC-6000 Plus
多機能オートサンプリングシステム

関連分野

＞ 食品・飲料

＞ ライフサイエンス

＞ 価格お問い合わせ

＞ 製品お問い合わせ

＞ 技術お問い合わせ

＞ その他お問い合わせ