機械学習による肥料の検査結果の予測

食料領域　上席主任研究官　川崎　賢太郎

肥料は農業の生産性の向上に不可欠ですが、外観からその品質を見極めることが困難です。こうした状態のことをミクロ経済学の用語では「情報の非対称性」と呼びますが、ノーベル経済学賞を受賞した著名な研究者アカロフは、情報の非対称性の下では、品質の粗悪な商品が市場を席巻してしまう可能性を指摘しました。

品質が表記どおりではない不正肥料の存在は、実際に途上国や明治期の日本などで確認されており（Bold et al, 2017、高橋, 2010）、我が国ではその対策として、1901年に肥料取締法を施行し、事前の登録制、立ち入り検査、違反に対する罰則などの導入を行いました。

しかし、この問題は現代においても解消されたとは言えない状況です。国内に約3,000社ある肥料業者のうち、約400社、のべ約700銘柄の肥料に対して毎年検査が行われていますが、近年ではその約２割で保証票の記載内容の誤り、保証成分量の不足及び有害成分の超過等の違反が見つかっています。そこで本研究では、AI（人工知能）の一種である機械学習（Machine Learning）を用いて検査結果の予測を行い、違反銘柄を一早く発見するための方策を検討しました。

１．機械学習とは

機械学習とは予測を行うための手法の総称です。本研究の目的は、検査に違反するか否かを、企業や肥料の特性といった様々な説明変数を用いて予測することですが、そこで鍵となるのは、説明変数をどのように選択するかです。説明変数の候補が多数ある場合、従来の方法（Logitモデルなどの回帰分析）では、分析者がそのうちの一部を主観的に選択する必要がありましたが、機械学習では説明変数が多数あっても、最適な組み合わせを客観的な方法で選択し、予測の精度を改善することができます。機械学習には様々なアルゴリズムがありますが、今回はランダムフォレスト（RF）とLASSOモデルを利用しました。

モデルの精度は適合率と再現率によって判定します。適合率とは「モデルで違反と予測された銘柄のうち、実際に違反している割合」のことで、第１表のTP/PMで計算されます。一方、再現率は、同TP/P0で計算され、「実際に違反している銘柄のうち、モデルで違反と予測された割合」を表します。

第１表　予測精度の評価

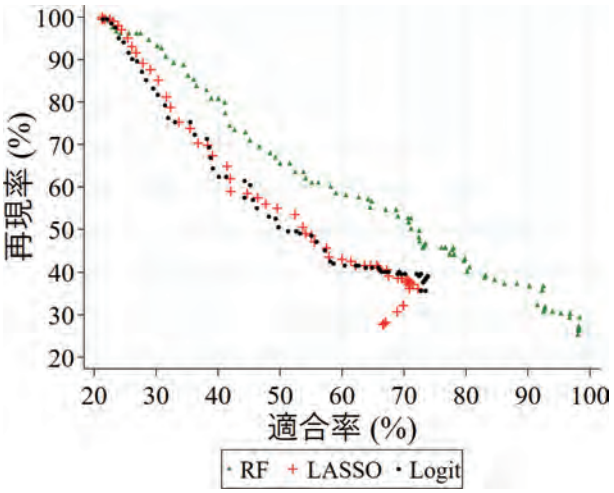
		モデルによる予測結果		計
		合格	違反	
実際の検査結果	合格	TN (True Negative)	FP (False Positive: 偽陽性)	NO
	違反	FN (False Negative: 偽陰性)	TP (True Positive)	PO
計		NM	PM	

資料：筆者作成

２．分析結果

データは2015年12月から2020年2月までの約3,000件の検査結果です。このうちランダムに選んだ約2,000件のデータを使って機械学習モデルを推計し、その結果を残る約1,000件に当てはめて精度を評価しました（第１図）。説明変数には企業や肥料の情報、前回の検査結果など様々な変数を使いました。なお同期間の検査における違反銘柄の割合は約21%です。

適合率と再現率が共に高い状態、すなわち図の右上方向にあるモデルほど、精度が高いことになります。したがって同図から、RFモデルが最もパフォーマンスが高く、次いでLASSOモデル、そして最も低いのが、機械学習ではない従来の手法であるLogitモデルであると言えます。



第１図　モデルの精度比較

資料：筆者作成

なお、同図では一つのモデルにつき99通りの結果が示されています。モデルからは個々の銘柄の違反「確率」が計算されるため、本稿では違反確率がV%以上であれば違反、V %未満であれば合格と判断することとし、この閾値Vとして１%から99%までの１%刻みの値を当てはめたためです。

RFモデルに関しても99通りの結果が示されていますが、この中でどの状態を目指すべきかについては、検査の効率性や公平性の観点から判断されるべきであり、次節で詳しく論じることとしますが、一例として適合率を現行の約２倍（41%）に引き上げることのできるRFモデルの詳細を第２表に示しました。

このモデルでは全銘柄の約４割（379/948）が違反と予測され、そのうち41%（156/379）が実際に違反しています。これが適合率です。また実際に違反している銘柄202件中、77%（156件）がモデルでも違反と予測されています。これが再現率です。逆に言えば、残る23%（46件）は、実際には違反しているものの、モデルでは違反なしと見逃されてしまうことになります。偽陰性と呼ばれる現象です。

３．検査業務への応用

偽陰性を防ぎながら、機械学習を実際の検査に応用する仕組みを考察したいと思います。

抜き打ち検査という特性上、検査対象の選定方法は公にされていませんが、まず機械学習を使わないベンチマークとして、全ての銘柄を順番に検査するケースを考えます。全銘柄数をN、毎年の検査数をＩ件とすれば、理論上はN/I年で全ての銘柄が検査され、全ての違反銘柄が摘発されることになります。この一巡に要する期間を「検査サイクル」と呼ぶことにします。我々の目標はこの検査サイクルを短縮することです。

次に、機械学習を導入するケースです。上述の通り、全ての検査を機械学習に頼ると、偽陰性、つまり実際には違反をしているにもかかわらず見逃されてしまう銘柄が発生してしまいます。これは公平性の観点から望ましくありません。そこで機械学習と順番制を組み合わせる仕組みを考えます。つまりRFによって違反と予測された銘柄（陽性グループと呼びます）から毎年θ×Ｉ件を順番に検査するのと並行して、RFによって違反なしと予測された銘柄（陰性グループと呼びます）についても、毎年（１－θ）×Ｉ件ずつ順番に検査することにします。θは陽性グループの検査割合です。こうした仕組みであれば、全ての銘柄がいずれは検査を受けますので、違反銘柄が見逃されることはありません。

第２表　RFモデルによる予測結果
（適合率41%のケース）

		予測結果		計
		合格	違反	
実際の検査結果	合格	523	223	746
	違反	46	156	202
計		569	379	948

資料：筆者作成

若干の計算を行うと、最も検査サイクルを短縮できるのは、適合率と再現率が60%前後のRFモデルを、θ=33%で利用した場合であることが分かりました。このモデルでは全銘柄のうち17%が違反ありの陽性グループ、残る83%は違反なしの陰性グループと分類されます。陽性グループは毎年0.33×I件ずつ検査されるため、検査サイクルはベンチマークの51%（=0.17N/0.33I）に短縮されます。陰性グループは毎年0.67×Ｉ件ずつ検査されるため、検査サイクルはベンチマークの1.24倍（=0.83N/0.67I）とやや長期化します。しかし両グループの検査サイクルを違反数で加重平均すると、検査サイクルはベンチマークに比べて約16%短縮されることになります。

つまり、機械学習の導入によって一部（陰性グループ）の検査が若干先延ばしされるものの、違反銘柄が多く含まれている陽性グループが頻繁に検査を受けることによって、全体で見れば検査サイクルが短縮され、違法な肥料が市場に出回るリスクを削減できることになります。なお、検査サイクルの短縮率は、銘柄数Nや検査数Ｉには一切依存しません。

さて、上記の仕組みでは銘柄を陽性・陰性の二つのグループに分割しましたが、機械学習では銘柄ごとに違反確率が計算されますので、この確率に基づいて更に細かくグループ分けをしたところ、検査サイクルは20%以上短縮されました。また業者ごとの生産量も考慮して検査順序を最適化すると、平均的な検査サイクルを50%以上短縮できることも分かりました。

今後は検査サイクルを一層短縮できる方法を検討し、実際の検査業務での活用を目指していきたいと考えています。

参考文献

Bold, T., Kaizzi, K. C., Svensson, J., and Yanagizawa-Drott, D. (2017) Lemon technologies and adoption: Measurement, theory and evidence from agricultural markets in Uganda. *The Quarterly Journal of Economics* 132(3) 1055-1100.
高橋周(2010)「明治後半における不正肥料問題: 新規参入の信頼獲得と農事試験場」『社会経済史学』76(3) 427-442.