【課題1】

１）アルゴリズムの評価指標を正答率ではなくAUC値に設定する

精度（この場合は正答率）が99.9%とあるが、全出品物に占める禁止出品物の割合が極めて低い（もしくは高い）場合は、正しく評価できない。
（仮に禁止出品物の割合が0.01%であれば、全ての出品物を負例／禁止出品物以外と予測するアルゴリズムの精度は99.99%となる）
よって、アルゴリズムの精度を正答率ではなくAUC値によって評価する

２）モデル構築と検証の枠組みを定める

未知のデータに対するアルゴリズムの精度を正当に評価する為、モデル構築に使用するトレーニングデータと検証用データを明確に切り分ける。
トレーニングデータと検証用データの割合はデータ量や正例（禁止出品物）と負例（禁止出品物以外）の割合等に基づいて決定する。
また、モデル構築に使用するアルゴリズムにおいて最適なハイパーパラメータを探索する必要がある場合には別途、ホールドアウトのサンプルを用意し最終的なモデルの精度評価を行う必要がある。

３）ビジネス面でのインパクトを評価する

出品されている品物が禁止商材ではないかを審査する業務に本アルゴリズムを導入した場合のビジネス面でのインパクトをシミュレートする。
業務のフローは以下を想定。
①全ての出品商品に対してアルゴリズムで禁止商材である確率を算出
②確率が一定の閾値を超えた場合にマニュアルレビュー（人による審査）を実施

＜評価ポイント＞
・マニュアルレビューが必要となる件数・割合
・False Positive（問題のない商品をマニュアルレビューの対象にしてしまう割合）とFalse Negative（禁止商材をスルーしてしまう割合）の
　各々の件数・割合
についてマニュアルレビューのコスト（人件費）や禁止商材をスルーしてしまった場合の損失額を勘案し、実用化出来るレベルに達しているかを
評価する。

【課題2】

・データのサンプルが十分ではない

⇒データ量を増やす。データ量が簡単に増やせない場合はバギング等の手法を用いてモデル構築を行う。評価も交差検証法等の方法を使用することで、サンプリングの際に発生するバイアスを極力排除する

・モデル構築用データにターゲットとリークしている特徴量が含まれている

⇒分類に寄与している特徴量別の重要度を算出し、飛び抜けて重要度の高い特徴量がある場合はリークを疑う

・特徴量の粒度が高すぎる（商品名称等の具体的すぎる情報が入っている）

⇒具体的な商品名等の特徴量については過学習に陥りやすいので「商品名」を「商品カテゴリ」に変換する等して特徴量の粒度を下げる

・モデルのアルゴリズムに問題がある

⇒ニューラルネットワークのようなか過学習しやすいアルゴリズムを使用している場合は階層の数を減らしたり、ロジスティック回帰のように過学習に陥りにくいアルゴリズムに変更する

・ トレーニングデータに使用した期間と検証用データに使用した期間が違う場合、時点間で外部環境の変化（景気の変動、広告やキャンペーンの有無等）が発生している

⇒外部環境を表す数値（GDPや失業率、広告出稿額やキャンペーンの有無）を特徴量に加えてモデルを作ことで、環境変化がある場合、その影響を予測に反映する（ベースとなる水準を調整する効果がある）

・運が悪い（トレーニング用データと検証用データが偶然にも異質になった）

⇒交差検証法を使って検証を行う。
サンプリングに使用する乱数のシード値を変える。

【課題3】

（ディープラーニングが用いられている理由）
近年、ディープラーニングが用いられている理由は主に以下の３点である。

１）トレーニングに使用出来るデータ量の増加
ニューラルネットは非線形な回帰分析を多層に渡って行う（一つのモデルのアウトプットが別のモデルのインプットになる）為、理論上は特徴量とターゲットの間のいかなる関係も表現することが可能である反面、少量のデータで行うと過学習に陥りやすい。
ディープラーニングが成果を上げている分野の一つに画像識別が挙げられるが、スマートフォンやSNSの普及によって膨大な数のデジタル写真データが収集可能になったことが大きいと思われる。
同様にディープラーニングが注目を浴びる契機となったAlpha Goに代表される囲碁や将棋の分野では、過去の人間棋士の対局データだけではなく、コンピューターどおしを仮想的に対局させることによって人工的に生成したデータをニューラルネットの学習に使用することで精度を上げている（強化学習）。

２）技術面での幾つかのブレークスルー
ニューラルネットの学習の肝はバックプロパゲーションと呼ばれる予測値と実績値の誤差を順次下位のレイヤーに伝搬していく仕組みである。
（数学の専門家ではないため厳密に理解している訳ではないが）従来のニューラルネットワークの学習では多層のネットワークを組んでも誤差が下層に伝搬しにくいという問題があったようで、パラメータが学習前の初期値に左右され、精度が出ないと言われていた。近年、特徴量を変換する関数やパラメータ初期値の設定方法、ドロップアウト等の過学習を防ぐ仕組み等が研究が進んだことで精度の向上が図れた。

３）ディープラーニングを手軽に実施出来るオープンソースの充実
１０年前にディープラーニングを使ったモデル開発を行おうとすると、フルスクラッチでコーディングするか高価な商用ソフト（SASやSPSS）を使用する以外に手段はなかった。近年ではGoogleのTensorFlowに代表されるオープンソースの充実によってディープラーニングを利用する敷居が低くなった。それによって多くの研究結果やノウハウの共有が行われ好循環を生んでいる。

上記を踏まえて個人的な見解は以下の通り。

１）ロジスティック回帰やランダム・フォレストでも十分な精度が出る
モデル構築対象となるビジネス分野のドメイン知識が豊富にある場合、特徴量の設定を工夫・洗練させることによって実用に十分な精度のモデルが作れることが多い

２）素人が安易に手を出すべき道具ではない
古くからニューラルネットワークが実用に用いられている分野としてクレジットカード取引の不正検知の分野が挙げられる。
この分野の専門家で元FICO社のアナリストと意見交換したことがあるが彼曰く、その道一筋２０年という米国本社のニューラルネットワーク職人がモデルを作っているらしく、彼はデータを見れば最適なネットワークの階層構造が”見える”とのこと。
この話を聞いて、ニューラルネットワークは素人が安易に手を出すべき道具ではないと感じ、距離を置いている。