

Sky Grid合同会社 AIシステム開発解説書 (CONFIDENTIAL)

山本洋佑 @ Sky Grid合同会社

Sky Grid合同会社 AIシステム開発解説書 (CONFIDENTIAL)

1. 執筆

- 2018年9月30日 Sky Grid合同会社 代表社員/AIエンジニア 山本洋佑 著

2. AIシステム開発工程

弊社Sky Grid合同会社では、AIを使ったシステム開発では下記フローにしたがってシステムを開発します。

1. PoC (Proof of Concept)
2. AIニューラルネットワーク開発
3. AI学習済みモデル開発
4. AI組み込みシステム開発

まずはPoCというAI開発の前処理となる検証フェーズを経て学習させるサンプル（以下、学習データ）を検証します。この検証を経て、最適となるAIニューラルネットワークを開発し、そのニューラルネットワークを使って学習済みモデルを開発します。これら3つの開発フェーズを経てAIを開発し、フェーズ4でAIを組み込むWebシステムやアプリ開発を行います。

要件によって大きく異なりますが、大まかなの工期、および費用感は以下の通りです。

項目	工期 (ヶ月)	費用 (万円)
PoC	1-3	50-300
ニューラルネットワーク開発	1-3	100-450
学習済みモデル開発	1-6	100-900
AI組み込みシステム開発	1-6	100-1000

これはあくまで目安となります。幸運なケースでは、既存技術の転用が可能であったり、PoCフェーズでニューラルネットワーク開発が同時に完了することで、そのフェーズの工数が0（費用もゼロ）になることもあります。この可能性については筆者でさえもクライアント様からのご要望を伺った後、PoCフェーズで検証してみないことにはわかりません。次節ではAIの開発フェーズを各々解説します。

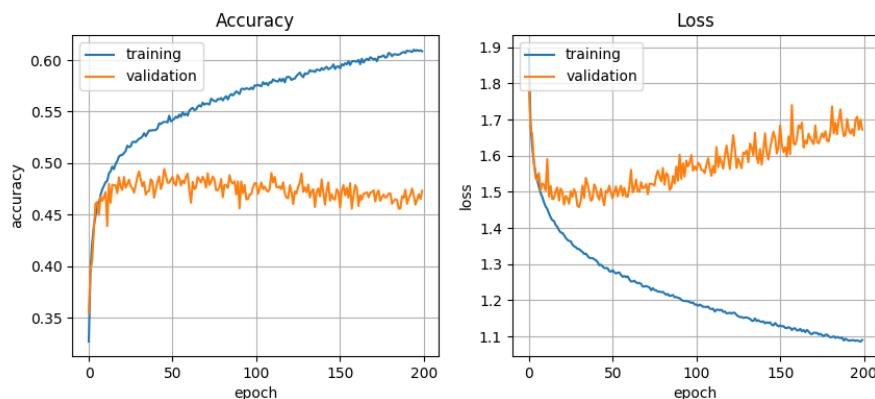
3. 検証・開発フェーズ解説

3.1. PoC (Proof of Concept)

AI開発では、学習データを使ってテスト・検証する以外に開発が可能か否か、加えてどの程度の精度を担保できるか断言することは不可能です。そこで弊社では、AI開発に着手する前に本検証フェーズでクライアント様から少量の学習データを頂いて、それらを検証させて頂きます。テストで検証する項目は以下の通りです。

1. どの程度の数量の学習データが必要か。
2. どのニューラルネットワーク（AIの種類）を選択するか。
3. ハイパーパラメータ（精度に関連するパラメータ）のチューニングにどれぐらいの時間が必要か。

少量の学習データから得た検証では、各ハイパーパラメータ毎に下図を描画し、考察します。

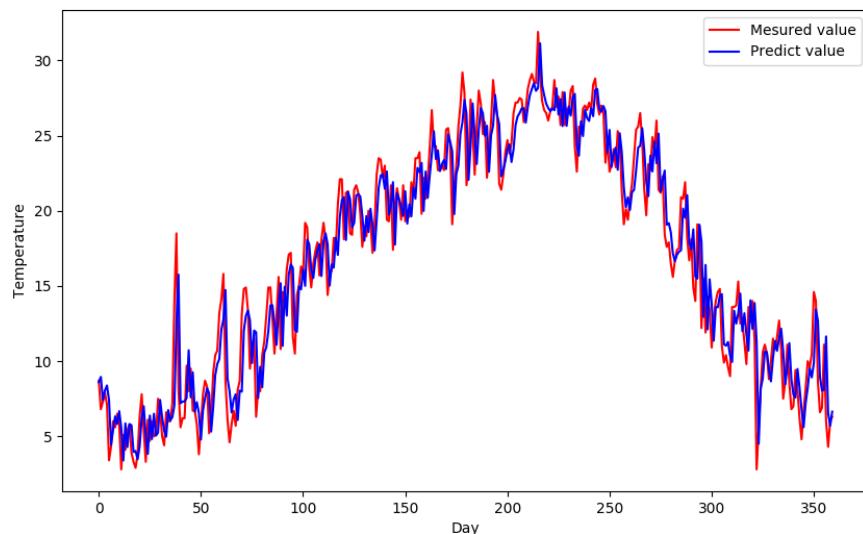


この図は筆者が開発したアヤメ種類分別AIシステムのPoCフェーズで描画したものとなります。両図共に横軸はepochというAIの学習回数を表し、今回の学習回数は200回となります。そして左図の縦軸は精度を表し、青線が1に近づくほど精度が高くなることを意味します。対して右図の縦軸は損失（間違い度合い）を表し、青線が0に近づくほど誤認が少なくなることを意味します。左図に着目すると学習回数200回の時点で精度が6割を超えていることから、要件によってはある程度実用レベルのAIが開発できていると判断できます。またこの青線は、ある一定の値に漸近しているわけではないので、学習回数を増やせば、さらに精度を上げられると推測できます。もし仮に青線が0.5などの一定の値に漸近し、その傾きがほとんど0となつた場合、それ以上学習回数を増やしても精度が上がらないことを示唆します。したがってニューラルネットワークの構成、あるいは他のハイパーパラメータのチューニング見直しが必要となります。

以上よりPoCフェーズを設けることで、行き当たりばったりなAI開発で生じる戻り作業のコストを事前に回避できます。弊社ではこのPoCの開発フェーズを設けることで、完成しうるAIの種類・精度を予測し、また必要となる学習データ数を推測します。したがってこの開発フェーズ完了後に、以降のニューラルネットワーク開発、および学習済みモデル開発フェーズを再見積もりさせて頂きます。クライアント様はその再見積もりを参考にして頂いて、AIの開発続行をご判断下さい。

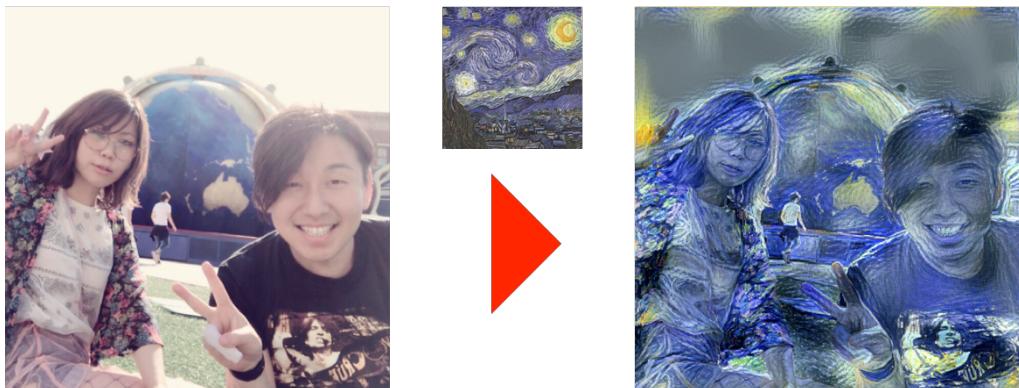
3.2. AIニューラルネットワーク開発

弊社ではクライアント様のご要望と学習データの形式に応じたニューラルネットワークを構築します。たとえば株価や気温の変遷を予測したい場合では、再帰的ニューラルネットワークのRNN (Recurrent Neural Network) や LSTM (Long short-term memory) を使います。筆者が開発したAIでLSTMを使って気温を予測した結果が下図になります。



横軸が日（365日）、縦軸が日毎の平均気温を示し、赤線が実測値、青線がAIによる予測値になります。10年間の平均気温変化をAIに学習させ、今日の気温から翌日の気温を予測しています。青線が1日ずつ先に描かれていくイメージです。予測値と実測値が比較的よくシンクロしていることが観察されます。

また複数のニューラルネットワークを使って、AIに絵を描かせることも可能です。下図は贋作者ネットワークと鑑定者ネットワークを作成し、2つのAIを競わせながら描いた図です。左図は筆者の写真ですが、AIにゴッホの写真を学習させ、右図を描かせています。



このように要件によって開発するニューラルネットワークは異なり、AIの挙動も異なります。またAIによる処理時間も大きく異なります。前者の天気の予測であれば、筆者のノートパソコンでも数十秒の処理時間で完了します。対して後者のゴッホの写真では、21時間ほどの処理時間を要します。これほどの処理時間をしていては、業務システムなどへの適用は難しいでしょう。

そこでニューラルネットワーク開発の完了後、弊社では学習済みモデルを開発します。学習済みモデルを開発した場合、ゴッホの写真AIでもそのモデルを使えば処理時間を数秒に短縮することができ、ほぼリアルタイムでの処理が可能となります。

3.3. AI学習済みモデル開発

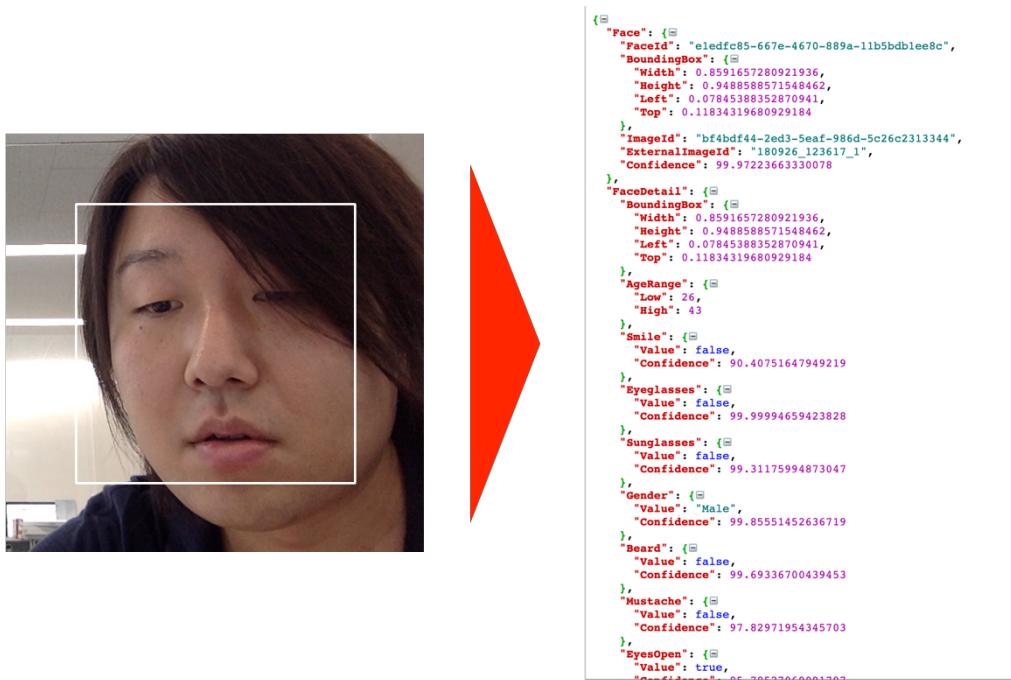
先述したように、各処理ごとにAIが学習を実行していくは、業務システムへの適用が困難です。そこで弊社は本開発フェーズで、実際に大量の学習データをニューラルネットワークに食わせ、学習済みのモデルを開発します。そのモデルを業務システムへ適用することでほぼリアルタイム処理が可能なAIシステムが開発されます。また大量の学習データをクライアント様からご用意いただけない場合、学習データを収集するシステムも同時に開発します。この有無によって本フェーズでの工数は大きく前後します。データがSNSなどから収集可能であればクローリング&スクレイピングシステムを開発します。対してWebなどから収集することができず、クライアント様が手入力しなければならないケースでは、学習モデル生成用の簡易システムやアプリを開発します。

このように本フェーズでは学習データの制作から学習済みモデルの開発までを行います。

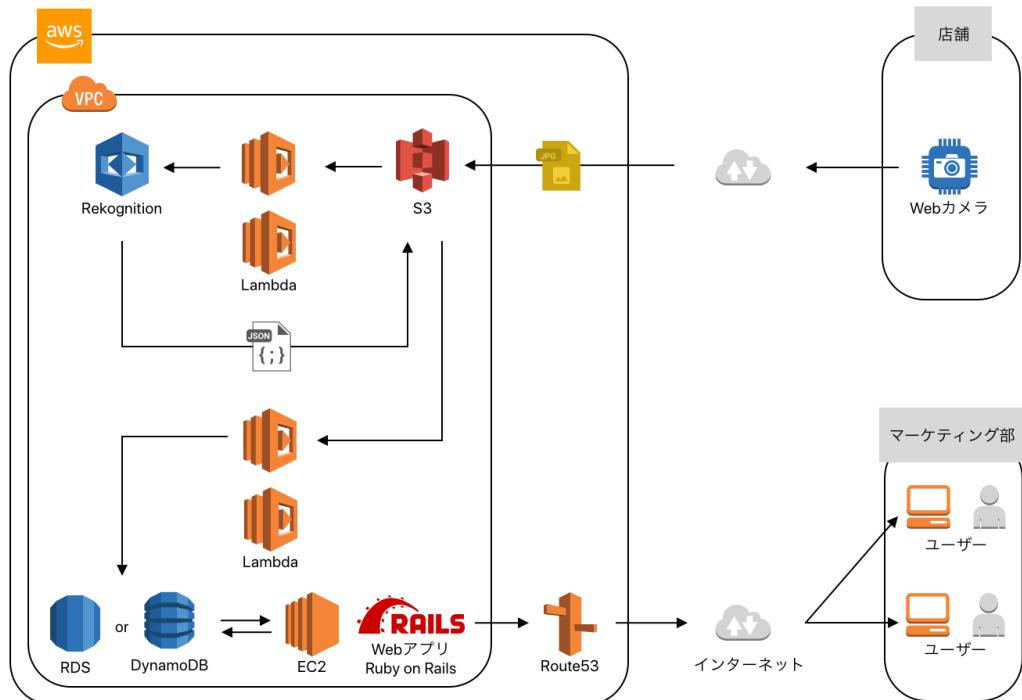
3.4. AI組み込みシステム開発

AIを実際に業務システムに組み込んだ場合、学習済みモデルを開発して処理時間を短縮することに加えて、AI処理を分散処理させるなどの工夫を盛り込みます。本節では弊社が独自に開発している顔認識AIマーケティングシステムを例に解説しま

す。そのシステムは店頭などにWebカメラを設置し、時間別による来店客層の調査、加えてリピート率調査などをするシステムとなっており、その稼働例は下図となります。



筆者の顔をWebカメラでリアルタイム撮影し、その属性をAIで分析しています。ここでは著者の性別や年齢、さらにはメガネの有無や笑顔度合いなど、様々なパラメータを正しく分析できていることがわかります。またこの処理を行うに際し、学習済みモデルを使ったAIで処理時間が数秒で済むとはいえ、人の出入りが激しい店頭では処理速度が追いつかない可能性があります。そこで弊社ではさらにAI処理をサーバー側で分散処理することでリアルタイム性を確保しています。具体的には、ある時間においてAIの処理が同時に複数必要となった場合（弊社システム例では1枚の写真に5人が映った場合など）、その処理数分AIのサーバーが自動で立ち上がり同時に処理をします。そして各々がその処理を完了すると自動で停止するといった仕組みを採用しています（下図参照）。



このようなシステム構成を設計することでリアルタイム性を保持しつつ、なおかつサーバー費用も節約することが可能となります。

4. 結言

この度はご一読頂き、誠にありがとうございます。AIというバズワードが飛び交う昨今、AI技術がより身近に、よりありふれたものになっていくことは疑いの余地もありません。本ドキュメントは弊社におけるAI技術を組み込んだシステム開発をフェーズ別で解説させて頂きました。AIは未だ発展途上であり、画像処理や再帰予測などの既存システムへのAI応用は科学者目線から見ても非常に有意義と感じます。お客様にとっても十分アピール商材となり得ることでしょうし、現場での実運用が十分現実的になってきていると感じています。お客様にはPoCフェーズの後、AI開発を見送り、従来のようなシステム開発を選択するといった選択肢もございます。まずは一度、気軽にご相談いただき、私共とご検討いただけますことを切に願っております。宜しくお願い申し上げます。

代表社員 山本洋佑