

2020年1月29日
株式会社インプレスR&D
<https://nextpublishing.jp/>

データ分析から AI アプリ開発まで優しく解説！
『Nuxt.js と Python で作る！ぬるさく AI アプリ開発入門』発行
技術の泉シリーズ、1月の新刊

インプレスグループで電子出版事業を手がける株式会社インプレス R&D は、『Nuxt.js と Python で作る！ぬるさく AI アプリ開発入門』（著者：深野 嗣、高山 和幸）を発行いたします。

最新の知見を発信する『技術の泉シリーズ』は、「技術書典」や「技術書同人誌博覧会」をはじめとした各種即売会や、勉強会・LT 会などで頒布された技術同人誌を底本とした商業書籍を刊行し、技術同人誌の普及と発展に貢献することを目指します。

『Nuxt.js と Python で作る！ぬるさく AI アプリ開発入門』
<https://nextpublishing.jp/isbn/9784844378129>



著者：深野 嗣、高山 和幸
小売希望価格：電子書籍版 2000 円（税別）／印刷書籍版 2400 円（税別）
電子書籍版フォーマット：EPUB3／Kindle Format8
印刷書籍版仕様：B5 判／カラー／本文 176 ページ
ISBN：978-4-8443-7812-9
発行：インプレス R&D

<<発行主旨・内容紹介>>

本書は、「モダンな AI アプリを開発する」ことを目標にした解説書です。AI アプリの開発という目的を明確にし、そこへの最短経路でデータ分析を行い、分析結果をモデル化し、Web アプリケーションに組み込むことによって、プロダクトを AI アプリに昇華させるためのチュートリアルです。

（本書は、次世代出版メソッド「NextPublishing」を使用し、出版されています。）

まず AI アプリ開発の進め方をステップを追って解説

図 1.1: AI開発可否判断シート

具体的には、次のようなステップで判断します。

1. 課題の有無
2. データの有無
3. システム化のニーズ
4. (人間 or 機械による) 判断の必要性
5. 投資対効果

本書では「目的」と表現していますが、ビジネスでAIを活用する場合は、解決したい「課題」が必要になるでしょう。課題が明確になったら、課題を解決できそうなデータが存在するか確認します。機械学習にはデータが必要不可欠だからです。

次に、解決したい課題にはシステムによる自動化ニーズがあるかどうかを確認してみましょう。ビジネス活用の場面では、実はユーザーは自動化を求めているというケースがちらほら見受けられます。ユーザーが求めているものを開発しても、使ってもらえませんよね。AI開発に活用できそうなデータを調べることに同じくらい、ユーザーがAIを求めているのかどうかを知ることは大切です。

そして、対象となる業務は人間による判断が求められている業務であるかどうか確認してください。求められているのであれば、AI開発を行うことで、人間による判断を機械による判断に置き換えられる可能性があります。もし、求められないのであれば、AI開発のかわりに、例えば最近話題のRPA（アール・ピー・エー）と呼ばれるシステムなどの他の手段を検討しましょう。

最後に、投資対効果も意識しましょう。AI開発による効果ができるだけ具体的に考えます。効果がみえない実装は開発のモチベーションも高まりません。また、ユーザーが喜んでくれば、AI開発を担う私たち自身も達成感を感じることが難いでしょう。

1.1.2 AIはヒューリスティックルールに強い

Heuristic Rule (ヒューリスティックルール)とは、必ずしも正しい答えを導けるわけではないが、ある程度のレベルで正解に近い答えを得ることができる規則のことです。完璧な精度ではない代わりに、答えが得られるまでにかかる時間が少ないという利点があります。例えば、決済シ

テムなどお金の授受を司るシステムは「大体、返金できた」なんていうことは絶対に起きてはいけません。機械学習を使ったAIアプリケーションは「大体でもいいから事前に判断してくれると助かる領域」に大きな効果を発揮します。機械学習は「一体どんな業務に役立つのか」と疑問に思う方は、ヒューリスティックルールがある領域を探してみてください。

1.1.3 目的に必要なデータセットを作る

最後に伝えたいことは、**データがなくてもAI開発をあきらめない**ことです。データがない場合は、データ収集からはじめましょう。実は、データがない段階からAI開発プロジェクトを始めたほうが、結果的に速く目的を達成するAI開発ができるケースが多いです。なぜなら、既存のデータを使ってAI開発を行う場合、データクレンジングに多くの時間が必要だからです。

AI開発に必要なのは、何よりもまず**目的変数を予測するために必要なデータセット**です。日本の大企業がどれだけビッグデータを持っていても、データ活用がうまく進まないのはいくつかの理由があります。その理由の1つは、蓄積されたデータがその目的のために集められたデータではないため、機械学習やAI開発に活用できないからです。

ただデータがあるだけでは、あなたの作りたいAIアプリを開発することはできません。これは紛れもない事実です。逆に、今はデータがなかったとしても、必要なデータを理解していれば、高品質なデータを効率よく収集でき、迅速にAIを開発することができるのです。

1.2 AIアプリ開発のステップを知ろう

AIを作るためには、1つずつステップを踏んでいく必要があります。AI開発に王道な道は、以下に、AI開発において積み上げるべきステップを明示しました。下から上、各ステップを順に積み上げていく必要があります。一つでも欠けると、最後のアプリケーション開発には到達できません。私はデータサイエンティストとして、クライアントがこの山登りのどのステップにいるのか、常に見極めることを大切にしています。

図 1.2: AI開発の7ステップ

| | |
|-------|-------------|
| STEP7 | アプリケーション開発 |
| STEP6 | 機械学習・モデリング |
| STEP5 | データ分析 |
| STEP4 | 学習用データセット作成 |
| STEP3 | 探索的データ解析 |
| STEP2 | データ収集・蓄積 |
| STEP1 | 目的設定 |

12 | 第1章 AIアプリ開発の進め方をしよう

第1章 AIアプリ開発の進め方をしよう | 13

AdobeXDをつかった最小の工数での UI 設計を解説

第4章 AdobeXDでアプリのUIを設計しよう

UIとは、ユーザーインターフェースのことであり、コンピュータとその機械の利用者との間の情報をやりとりするための接点を指します。モデルを組み込んだAIアプリ開発において、UI設計はとても大切です。

本書で開発するAIアプリとユーザーの接点は画面です。その画面からデータを取得し、AIは予測します。本章では、AdobeXDと呼ばれるモックアップツールで実際に開発するAIアプリの画面を最小の工数で設計します。

4.1 AIアプリ開発においてUI設計が大切な理由

本書で開発するAIアプリのコア機能である教師あり学習モデルは、目的変数を予測するための説明変数をれなく用意する必要があります。このとき、説明変数のデータ収集方法について考えてみましょう。

実際にアプリからユーザーのデータ収集する場合は、**明示的データ**と**暗黙的データ**について知っておくと良いでしょう。明示的データと暗黙的データはそれぞれ強みと弱みがあり、取得方法も異なります。

これを理解しないで進めてしまうと、既存データを分析し、データセットを作り、予測モデルを構築することができても、モデルをアップデートするための新しいデータは手に入りません。AIアプリは、常に新しいデータを収集できるUIになっている必要があります。

4.2 機械学習に必要なデータをいかに収集するか

4.2.1 明示的データ

明示的データとは**ユーザー自身が認識して入力したデータ**を指します。主にデモグラフィックデータなどの個人情報、コンテンツに対する主観的評価に関するデータがそれに当たります。デモグラフィックデータとは、年齢、性別、居住地、家族構成、職業など、人口統計学的なデータのことを指します。

明示的データのメリットはデータそのものの信頼性が高いことです。一方、デメリットはユーザーに入力する負担がかかることです。みなさんもECを利用するときやインターネットで銀行・証券口座を開設しようとするときに、会員情報入力画面を見て「面倒くさい」と思ったことがあるのではないのでしょうか。明示的データを必要とするモデルの場合は、ユーザーにとって負担にならないUI設計を考える必要があります。

実際に、明示的データを取得しているサービス例は次のとおりです。

- ・YouTubeの動画評価「Good or Bad」
- ・Amazonの商品評価「5つ星評価」
- ・各サービスの会員情報の入力
- ・ゲーティングアプリでの「スワイプ (好み or 好まない)」

4.2.2 暗黙的データ

暗黙的データとは**ユーザー自身が入力していることを意識していないデータ**を指します。蓄積された、Webサイト上での行動ログなどがそれにあたります。ビッグデータと呼ばれる量のデータの9割以上は、これらの暗黙的データのこの指していると言った間違いはないでしょう。例えば、Webサービスの画面上での「プレビュー」「クリック」「フォロー」「購入履歴」など、これらはユーザーはデータを与えていることを強く意識していません。ユーザーに聞かずして取得できているデータです。

暗黙的データのメリットは、ユーザーがデータを入力する必要がないためユーザーの負担が少ないことです。デメリットは、その分データの信頼性が低くなることです。そのため、ユーザーの行動について仮説を作り、相関関係と因果関係をしっかりと見極めてデータ分析を行い、仮説を検証する必要があります。

実際に、暗黙的データを取得しているサービス例は次のとおりです。

- ・Netflixの視聴履歴 (再生時間、視聴開始日時など)
- ・Webアプリ等のボタンクリック、表示履歴
- ・Twitterのフォロワーリスト、フォローワーリスト
- ・海宝閣 (オオバコ) での購入履歴
- ・SmartNewsのような機械学習によるニュースキュレーションアプリの購読履歴

4.2.3 機械学習にはどちらのデータが必要か

機械学習を活用しているアプリケーションの殆どは、明示的データと暗黙的データを両方使っています。GAF Aをはじめとする企業は暗黙データを大量に保有しているため、ユーザーの趣味嗜好に対して予測できます。一方、個人開発の場合は、行動ログのような暗黙的データを収集するデータベース設計が必要となり、開発の負担が大きくなります。そのため、本書で開発するアプリでは、**明示的データを収集し、それらを入力する画面を作成しましょう。**

4.3 AdobeXDでアプリのUIを設計しよう

繰り返しますが、AIアプリ開発においてUI設計が重要な理由は、**ユーザーにデータを入力してもらう必要がある**からです。モデルの予測に必要なデータをユーザーに入力してもらうために、開発者は色々な方法を考える必要があります。

1. 任意項目ではなく必須項目にする
2. 必須項目に計算不能値を入力できないようにする
3. ユーザーにとって負担にならない入力方法を設計する

今回開発するAIアプリは、必要なすべての説明変数が入力されれば、モデルは機能しませ

56 | 第4章 AdobeXDでアプリのUIを設計しよう

第4章 AdobeXDでアプリのUIを設計しよう | 57

Nuxt.js でアプリのフロントを、Python で API を作る過程を丁寧に紹介

第5章 Nuxt.jsでAIアプリのフロントを作ろう

それでは、今まで設計したAIアプリを実際に実装していきましょう。
ここからは、筆者の個人的な興味と偏見も若干含んだ技術スタックでモダンなAIアプリを実装していきます。
今回のサンプルアプリケーションのコードはすべて <https://github.com/pc02699/NullSuck-AI> で見ることができます。本文の説明と合わせて参照いただくと理解が進みます。
また、本アプリは <https://ai.pco2699.net> にデプロイされているので、興味がある方はアクセスしてみてください。

5.1 今回作るアプリ

次のような構成でアプリを実装していきます。

図5-1: アプリの構成図

```
graph LR; Frontend[Nuxt.js] --> API[API responder.py]; API --> DB[(MySQL)];
```

フロント: ワインの属性値を入力する
API: ワインの属性値を受け付けDB機械学習器のデータをもとに結果を返す
DB: ワインのマスターデータを保持

本アプリは主に、フロントをNuxt.jsというフロントエンドフレームワークを用いて実装していきます。本章では、そのフロント部の実装を中心に説明していきます。

5.1.1 全体のプロジェクト構成

全体のプロジェクト構成は次のとおりです。

```
NullSuck-AI
├── client
├── server
├── docker-compose.yml
├── k8s
├── k8s.yml
├── LICENSE
└── makefile
```

5 directories, 5 files

client
本章で説明するNuxt.jsを用いたフロントエンドのソースコードを格納しています。

server
第6章「PythonでAPIを作成しよう」で説明するPythonを用いたバックエンドのソースコードを格納しています。

その他
第7章「AIアプリをデプロイしよう」で説明するデプロイに用いるファイルなどを格納しています。

5.2 Nuxt.jsとは

Nuxt.jsとは、Vue.jsアプリケーションを作るためのフレームワークです。

5.2.1 Vue.js

Vue.jsはフロントエンド/JavaScriptのフレームワークです。JavaScript/HTML5/CSS3で、フロントエンドをガリガリ書こうとすると、レイアウトとロジックが混在します。そのため、どうしてもコードが汚くなったり、複雑な動きの実装が難しい状況でした。
そこで、登場したのがフロントエンドフレームワークです。フロントエンドフレームワークは、さまざまな手法で先の問題を解決することを試んでいます。近年だとAngular/React/Vue.jsが有名です。Vue.jsはその中でも、HTML5/CSS3などの知識を活かして学習コストが低いフレームワークとして近年人気が出ています。

<<目次>>

- 第1章 AI アプリ開発の進め方を知ろう
 - 1.1 目的から考えよう
 - 1.2 AI アプリ開発のステップを知ろう
- 第2章 Google Colaboratory で環境を構築しよう
 - 2.1 Google Colaboratory にアクセスする
 - 2.2 データ保存先として Google ドライブを利用する
 - 2.3 Markdown 記法を積極的に使おう
- 第3章 Scikit-learn でモデルを構築しよう
 - 3.1 ワインの美味しさを予測してみよう
 - 3.2 探索的データ解析
 - 3.3 Googleドライブをマウントする
 - 3.4 データ準備
 - 3.5 機械学習用データセット作成
 - 3.6 データ分析
 - 3.7 モデリング
- 第4章 AdobeXD でアプリの UI を設計しよう
 - 4.1 AI アプリ開発において UI 設計が大切な理由
 - 4.2 機械学習に必要なデータをいかに収集するか
 - 4.3 AdobeXD でアプリの UI を設計しよう
 - 4.4 UI 設計の手順
 - 4.5 AdobeXD を使って画面をデザインしよう

第5章 Nuxt.js でAIアプリのフロントを作ろう

5.1 今回作るアプリ

5.2 Nuxt.js とは

5.3 実際にプロジェクトを作ってみよう

5.4 アプリの構成をVuexifyで作ろう

5.5 APIに投げるためにVuexを実装しよう

第6章 PythonでAPIを作成しよう

6.1 今回、利用する構成

6.2 簡単なAPIを作ってresponderで動かしてみよう

6.3 プロジェクトの構成

6.4 ハンドラを実装しよう

6.5 DBと連携できるようにしよう

6.6 機械学習のモデルを組み込もう

6.7 クライアントとAPIを連結させてみよう

第7章 AIアプリをデプロイしよう

7.1 今回のデプロイ環境の構成

7.2 ローカル環境をDocker化しよう

7.3 DockerイメージをGoogle Container Registryに登録してみよう

7.4 本番環境をKubernetes・GKE・CloudSQLで作成しよう

7.5 ドメインの設定をしよう

7.6 CircleCIで自動でデプロイするようにしよう

<<著者紹介>>

深野 嗣(ふかの ひで)

北海道出身の1991年生まれ、28歳独身。必要なことを必要な分だけ伝えることが得意なデータサイエンティスト。2年前に貯金なしで独立し起業を志すも過労と栄養失調により気絶。Pythonとデータサイエンスのおかげで何とか貯金ができるようになり、最近再び独立。休日は、推薦システムの勉強のためという建前でNetflixをどっぷり見ちゃって罪悪感がすごい。ビギナーズラックがすごくて初出場・初体験の受賞経験だけが豊富。何かと2回目以降うまくいかないことが最近の悩み。運に実力が追いつくようにもっと頑張るンゴ！

高山 和幸(たかやま かずゆき)

東京都世田谷区在住の30歳、資産運用サービスを行うスタートアップに通勤しながら本を書くのが好きな社畜バックエンドエンジニア。HackDay2018無冠。2ヶ月準備してGoogleを受けるも、Phone Interviewであっさり落ちる。Googleのインタビューの準備で勉強したアルゴリズムが楽しく、技術書典8で「アルゴリズムミニアプリ」という本を出し、「これは絶対売れる」と確信するものの、在庫を半分余らせる大爆死を経験する。仕事と執筆が両立できず、帯状疱疹になるなど、30歳になってから体力の低下をひしひしと感じている。

<<販売ストア>>

電子書籍:

Amazon Kindleストア、楽天koboイーブックストア、Apple Books、紀伊國屋書店 Kinoppy、Google Play Store、honto 電子書籍ストア、Sony Reader Store、BookLive!、BOOK☆WALKER

印刷書籍:

Amazon.co.jp、三省堂書店オンデマンド、honto ネットストア、楽天ブックス

※ 各ストアでの販売は準備が整いしたい開始されます。

※ 全国の一般書店からもご注文いただけます。

【インプレス R&D】 <https://nextpublishing.jp/>

株式会社インプレス R&D(本社:東京都千代田区、代表取締役社長:井芹昌信)は、デジタルファーストの次世代型電子出版プラットフォーム「NextPublishing」を運営する企業です。また自らも、NextPublishing を使った「インターネット白書」の出版など IT 関連メディア事業を展開しています。

※NextPublishing は、インプレス R&D が開発した電子出版プラットフォーム(またはメソッド)の名称です。電子書籍と印刷書籍の同時制作、プリント・オンデマンド(POD)による品切れ解消などの伝統的出版の課題を解決しています。これにより、伝統的出版では経済的に困難な多品種少部数の出版を可能にし、優秀な個人や組織が持つ多様な知の流通を目指しています。

【インプレスグループ】 <https://www.impressholdings.com/>

株式会社インプレスホールディングス(本社:東京都千代田区、代表取締役:唐島夏生、証券コード:東証1部9479)を持株会社とするメディアグループ。「IT」「音楽」「デザイン」「山岳・自然」「旅・鉄道」「学術・理工学」を主要テーマに専門性の高いメディア&サービスおよびソリューション事業を展開しています。さらに、コンテンツビジネスのプラットフォーム開発・運営も手がけています。

【お問い合わせ先】

株式会社インプレス R&D NextPublishing センター

TEL 03-6837-4820

電子メール: np-info@impress.co.jp