

スキルシート

ふりがな		みながわ れんたろう		年齢	性別	最寄駅
氏名		皆川 連太郎		27歳	男性	京王線 仙川駅
OS		Windows				
言語		Python, C++				
ツール		PyTorch				
モデル		GBDT、NN、BERT、Transformer、UNet、CNN、ViT、DinoV2、LightGlue				
自己PR		【課題発見～改善の実行力】 ・モデルの性能が悪化する技術課題(RAWとRGBの信号レベルに乖離がある)に対して、モデル内部の切り出しやデータ解析など様々なアプローチで原因解明に努めた。 ・学習データに信号レベル方向での不足があると仮定することで、データ拡張の施策を考案できた。 ・ホワイトバランス補正値などを前提に設計し、モデル性能を最大化させる専門的なデータ拡張を実現。 【フルスタックな ML スキル】 ・データ分析・前処理・特徴量設計～モデル実装・評価まで単独でAI 開発のフローに対応可能。 ・Kaggle にて合計 10 回の上位入賞の実績があり、複数分野・モデルの機械学習タスクを 3 年ほど経験。 ・経験分野は画像(3D 画像再構成・患部画像分析)、自然言語(Image2Seq・AI文章分類)、テーブルデータ(生存時間分析・ゲームAI勝率予測)、シミュレーション(軽量チェスエンジン開発)など。				
項番	期間		業務内容 ※項番1～3はすべて同じ企業内の案件です ※項番4～9は個人開発です	担当業務	環境・言語等	役割
1	2023年 12月	20ヵ月間	【外資系通信企業】 情報通信技術(ICT)およびスマートデバイス開発 【プロジェクト概要】 既存のデモザイクモデルの技術課題分析 【実績・取り組み】 ・RAW画像(カメラ内部の生画像)からRGB画像(一般的な画像)に変換するアルゴリズム(デモザイク)に対し、機械学習を用いたモデル開発・改善を行う。 ・既存モデルでは暗所領域において性能が悪化する技術課題があったため、以下のようなアプローチで問題の原因を解析する。 ・モデル内部を切り出すことでモデル特徴量の傾向を把握する。 ・技術論文を参照して、暗部領域に強みを持つモデル構造を研究する。 ・学習データの内部分析、あるいはデータ生成過程の見直し。	カメラ AI アルゴリズム開発	【言語】 Python 【OS】 Windows 【フレームワーク】 PyTorch	メンバー
	2025年 7月					
2	2023年 12月	20ヵ月間	【外資系通信企業】 情報通信技術(ICT)およびスマートデバイス開発 【プロジェクト概要】 デモザイクモデルの課題解決のための施策立案 【実績・取り組み】 ・技術課題分析を基に以下のようなアプローチで施策を考案・実装し性能比較を実験した。 ・モデル内部のResNetブロックに対し、パラメータや構成要素などを変更する。 ・モデルの入力・出力部分においてRAW/RGB形式の変更、追加。 ・学習データ生成過程でノイズ付与やクラスターごとに仕分けなど。 ・モデル学習時のパラメータ・パイプラインの変更。 ・モデル出力に対する線形変換などルールベースの後処理。 ・推論環境の見直し。	カメラ AI アルゴリズム開発	【言語】 Python 【OS】 Windows 【フレームワーク】 PyTorch	メンバー
	2025年 7月					
3	2023年 12月	20ヵ月間	【外資系通信企業】 情報通信技術(ICT)およびスマートデバイス開発 【プロジェクト概要】 デモザイクに関する新たな機械学習モデルの実装・開発 【実績・取り組み】 ・性能比較を基に以下のアプローチで技術課題を解決した機械学習モデルを開発。 ・学習データを一律に暗くすると、暗所領域での性能が改善したのが発見のきっかけ。 ・上記発見から学習データ自体に不足はないが(通常時は精度良好のため)、偏りがあると仮定し、データ拡張により不足を補う。 ・学習データの生成過程で補正値やビット数変換など特定の処理が入るため、それらを前提にデータ拡張を設計する。 ・一様分布ではなくガウス分布などを用いて、元のデータ中心に沿った信号レベル分布を形成する。 ・以上のデータ拡張をモデル学習時のパイプラインに追加し、技術課題を解決するに足る精度改善(通常領域と遜色ない精度)を達成した新モデルを開発した。	カメラ AI アルゴリズム開発	【言語】 Python 【OS】 Windows 【フレームワーク】 PyTorch	メンバー
	2025年 7月					

4	2024年 9月	3ヵ月間	<p>【個人開発】 予測モデリング及び分析手法関連プラットフォーム</p> <p>【プロジェクト概要】 モンテカルロ木探索(MCTS)AIエージェントのボードゲーム勝率予測</p> <p>【実績・取り組み】 ・1000種類の世界各地のボードゲームごとのMCTSエージェント同士の対戦結果を予測。 ・情報量の少ない特徴量を削除、ゲーム説明の文章データに対してWord2Vec等の特徴量作成。 ・クラス不均衡に対応させたCV戦略を採用。 ・二段階モデリングにより、機械学習モデルの多様化・精度向上を実現。 ・20位 (1608チーム中) にて銀メダル獲得。</p>	テーブルデータ AI アルゴリズム開発	<p>【技術】 Word2Vec StratifiedGroup Kfold LightGBM XGBoost CatBoost</p>	メンバー
	2024年 11月					
5	2024年 3月	3ヵ月間	<p>【コンペティション出場】 予測モデリング及び分析手法関連プラットフォーム</p> <p>【プロジェクト概要】 画像マッチング技術を用いた3次元再構成</p> <p>【実績・取り組み】 ・複数の2次元画像シーンに対し、3次元空間に再構成・投影を行う。 ・一連のパイプラインは画像ペア選定、特徴点抽出、マッチング、3次元復元の順に実装する。 ・最初に2枚1組のグローバル特徴量の近い画像ペアに限定することで、全体処理時間を削減。 ・ALIKEDなどのモデルを用いて画像から特徴点を抽出する。ハッシュ保存によって処理時間を軽減。 ・回転やクロップなどを用いることで、より多くの特徴点を抽出でき、精度向上に繋がる。 ・得られた特徴点を基に画像ペアごとに点マッチングを行う。(LightGlueなど) ・COLMAPを用いて、2次元マッチング情報を基に3次元空間にマッピングする。 ・31位 (929チーム中) にて銀メダル獲得。</p>	カメラ AI アルゴリズム開発	<p>【技術】 DinoV2 OpenCV ALIKED LightGlue COLMAP</p>	メンバー
	2024年 5月					
6	2023年 11月	3ヵ月間	<p>【コンペティション出場】 予測モデリング及び分析手法関連プラットフォーム</p> <p>【プロジェクト概要】 大学レポートにおけるLLM生成文章の分類</p> <p>【実績・取り組み】 ・大学の英語レポートに対して、学生の文章か生成AIの文章のどちらかを分類する。 ・各文章に対して、Tokenizerを用いてトークン化する。 ・トークンデータに対して、Vectorizerを用いてTf-idfベクトル表現を抽出する。 ・GBDT系モデルを中心に、ベクトルデータを入力として分類タスクを解く。 ・108位 (4358チーム中) にて銀メダル獲得。</p>	自然言語 AI アルゴリズム開発	<p>【技術】 Transformer DeBERTaV2 LightGBM</p>	メンバー
	2024年 1月					
7	2024年 12月	3ヵ月間	<p>【コンペティション出場】 予測モデリング及び分析手法関連プラットフォーム</p> <p>【プロジェクト概要】 限定的な計算資源で動作する軽量チェスエンジン開発</p> <p>【実績・取り組み】 ・5MBのRAM、ファイルサイズが最大64KBの限られた計算資源で動作するチェスエンジンの開発。 ・軽量性と強さの観点からベースコードにEthereal ver11.75を採用。 ・5MB以下のメモリ領域に収めるため、ハッシュテーブルのサイズ変更。 ・ファイルサイズを抑えるため、コンパイル時に-Os最適化オプション使用、tar圧縮。 ・51位 (1120チーム中) にて銀メダル獲得。</p>	シミュレーション AI アルゴリズム開発	<p>【技術】 C++ UCI</p>	メンバー
	2025年 2月					
8	2023年 3月	3ヵ月間	<p>【コンペティション出場】 予測モデリング及び分析手法関連プラットフォーム</p> <p>【プロジェクト概要】 生成AI画像の元の文章を逆予測</p> <p>【実績・取り組み】 ・StableDiffusionで生成された画像に対して、生成元になった文章を逆予測するImage2Seqタスク。 ・モデルは大別して2つ、1つはCLIPやCoCaなどのキャプションモデルをゼロショット学習で用いる。 ・もう1つはViT(Vision Transformer)を用いる。正解文章とのコサイン類似度を損失関数に設定してモデルを学習。 ・計算資源の制約上、モデルのサイズを適度に落としつつモデル作成、最後にアンサンブルする。 ・43位 (1231チーム中) にて銀メダル獲得。</p>	画像・自然言語 AI アルゴリズム開発	<p>【技術】 StableDiffusion CLIP CoCa VisionTransformer</p>	メンバー
	2023年 5月					

9	2025年 6月	3ヵ月間	<p>【コンペティション出場】 予測モデリング及び分析手法関連プラットフォーム</p> <p>【プロジェクト概要】 手首装着ウェアラブルセンサーから患者の反復行動分類</p> <p>【実績・取り組み】 ・時間・体温・動きに関するセンサーデータから、髪を引っ張る・皮膚を掻くなどの特定の反復行動を検知・分類する。 ・時間データはパディング戦略を最適化、体温データは接地面積ごとの統計量追加、動きデータは角度情報作成などデータ前処理を行う。 ・上記データ群を各自のResNetSEBlockに入れて中間特徴量を獲得。 ・各セクションの中間特徴量を患者の行動インデックスごとに整理し、時間特徴量としてLSTM・GRUのRNN系、またはRoBERTa・DeBERTaのBERT系モデルを用いて学習。 ・154位(2657チーム中)にて銅メダル獲得。</p>	テーブルデータ AI アルゴリズム開発	<p>【技術】 ResNet LSTM GRU RoBERTa DeBERTa</p>	メンバー
	2025年 8月					
10	2025年 7月	3ヵ月間	<p>【コンペティション出場】 予測モデリング及び分析手法関連プラットフォーム</p> <p>【プロジェクト概要】 高分子化学ポリマー構造を用いた指標予測</p> <p>【実績・取り組み】 ・ポリマー構造を記述する文字列であるSMILESデータを用いて、密度、熱伝導率、ガラス転移温度、回転半径、自由体積を予測する。 ・オープンソースデータを用いたデータ拡張、SMILESに対する重複削除やランダム並び替え、ノイズ付与などの前処理。 ・RDKit・Mordredなどの分子記述子を用いて数百個の特徴量作成、各予測対象ごとに分子特徴量を厳選。 ・分子特徴量を入力としてGDBT・NN・GNNなどのテーブル用AIモデル用いて学習、SMILES文字列を入力としてChemBERTaなどのNLPモデルを学習し、出力結果をアンサンブル。 ・187位(2240チーム中)にて銅メダル獲得。</p>	自然言語 AI アルゴリズム開発	<p>【技術】 RDKit Mordred LightGBM NN ChemBERTa</p>	メンバー
	2025年 9月					