Owen Wienczkowski – AI/ML Project Portfolio

GitHub: https://github.com/owenwienczkowski
LinkedIn: https://www.linkedin.com/in/owenwienczkowski
Email: owen.wienczkowski@gmail.com
Location: Denver, CO

This portfolio presents five projects selected from my GitHub that demonstrate key skills in machine learning, NLP, transformer-based modeling, deployment, and statistical analysis. Each project showcases a different capability relevant to real-world problem solving and industry-aligned ML engineering practices. The descriptions below highlight the problem context, applied methods, tools, and relevant skills. These projects cover a range of modern AI applications, from transformer-based sentiment classification to real-time QA with LLMs. Each is structured for clarity, modularity, and practical reuse.

Project Summaries:

Prompt Engineering Toolkit for Gemini QA

Project Motivation:

LLMs often require domain adaptation for document-based question answering. This tool addresses the challenge of prompting Gemini to reason effectively over uploaded documents.

Solution Overview:

Built a local QA interface that applies zero-shot, few-shot, and chain-of-thought prompts to Gemini, enabling comparative testing across arbitrary document types. Modular design allows flexible experimentation.

Skills:

- Engineered zero-shot, few-shot, and CoT prompt strategies for Gemini
- Integrated Google Gemini API for secure document-based QA
- Designed modular Python architecture for prompt reuse and extensibility
- Achieved consistent QA performance on multi-topic documents with varying prompt types
- Focused on reasoning via prompting without model fine-tuning

Repo: https://github.com/owenwienczkowski/Prompt-Based-Gemini-QA-Assistant

Structured Credit Risk Classifier with Evaluation Pipeline

Project Motivation:

Classifying credit risk using structured data is crucial in financial decision-making. This project models such a system using real-world features in fully modular architecture.

Solution Overview:

Built a modular ML pipeline to classify borrower risk using logistic regression, random forest, and gradient boosting. Included robust preprocessing, outlier handling, and manual feature reduction. Evaluated models on accuracy, F1-score, AUC, and precision-recall curve.

Skills:

- Built fully modular classification pipeline for structured risk modeling
- Implemented logistic regression, random forest, and gradient boosting models
- Engineered feature preprocessing with outlier detection, one-hot encoding, and imputation
- Achieved AUC of 0.95 and balanced F1-scores on held-out evaluation data
- Visualized results using PR curves and classification thresholds

Repo: https://github.com/owenwienczkowski/credit-risk-text-classifier

Fine-Tuned IMDb Sentiment Classifier with Hugging Face

Project Motivation:

Transformers offer strong baseline performance for text sentiment classification. This project explores fine-tuning Hugging Face models for structured inference.

Solution Overview:

Fine-tuned DistilBERT and Tabularisai for sentiment analysis on IMDb data. Implemented tokenization, decoding, label bucketing, and classification report evaluation.

Skills:

- Fine-tuned DistilBERT and Tabularisai on IMDb reviews (binary + bucketing)
- Tokenized and mapped labels to standard output classes
- Customized classification reporting for precision/recall at class level
- Reached >91% test accuracy on balanced IMDb subset
- Designed modular training/evaluation script using Hugging Face Transformers

Repo: https://github.com/owenwienczkowski/Hugging-Face-IMDb-Fine-Tuned-Sentiment-Classifier

Deployed Sentiment Classifier with Binary + Multiclass Modes

Project Motivation:

Many real-world applications require deploying sentiment models that support multiple output granularities. This tool provides a deployed interface for binary and multiclass text classification.

Solution Overview:

Built a Streamlit tool offering binary and multiclass classification modes using pretrained Hugging Face models. Integrated modular evaluation and user-facing controls to make the system extensible and usable without technical expertise

Skills:

- Developed real-time sentiment analysis interface using Streamlit
- Deployed binary/multiclass inference pipeline using Hugging Face Transformers
- Enabled custom class toggling and review-level output granularity
- Achieved consistent user-facing performance across categories
- Designed extensible codebase for non-technical use and model swapping

Repo: https://github.com/owenwienczkowski/sentiment-analysis-huggingface

UKContentment Predictor

Project Motivation:

Understanding predictors of life satisfaction is a key challenge in national well-being policy analysis. This project builds a multiple linear regression model using UK contentment data.

Solution Overview:

Applied regression modeling and diagnostics using R to predict contentment scores. Included residual analysis, outlier detection, and variable selection with strong visualization support.

Skills:

- Conducted multiple linear regression on UK census-derived satisfaction dataset
- Cleaned, filtered, and visualized residuals and influential points
- Created interpretable model using standardized coefficients
- Generated statistical summaries and variable diagnostics in R
- Delivered explanatory plots to support policy-facing insights

Repo: https://github.com/owenwienczkowski/UKContentmentPredictor