# Final Year Project Ideas

Justin Lim

May 4, 2025

## Contents

# 1 Bias Detection in Social Media News using NLP

**Target Audience**

Tertiary students and young adult readers who frequently consume news through social media and may be vulnerable to biased or emotionally manipulative content. This project aims to enhance their media literacy and critical reading skills.

**Template Chosen**

12.1 Project Idea 1: Identifying research methodologies that are used in research in the computing disciplines

**Aim**

To develop an NLP-based tool that helps students detect emotionally charged or misleading language in social media news snippets, enabling them to recognize biased framing and misinformation indicators.

**Objectives**

1. Develop a functional prototype that detects and highlights bias-indicative language (e.g., emotionally charged phrases, vague quantifiers) in social media-style news content, with at least 75% precision on a curated test set.

2. Demonstrate that users of the tool are able to identify biased statements more accurately, with at least a 20% improvement in post-test scores over a pre-test baseline in a simulated user study.

3. Generate annotated outputs with visual cues (e.g., highlighted words, explanation tooltips) that users report as helpful in at least 70% of feedback responses collected via short usability surveys.

**Justifications**

- **Why:** In Singapore and globally, news is increasingly consumed via social media platforms where misleading headlines, emotional language, and framing bias are common. Young readers may lack the media literacy skills to spot these manipulations, making them more susceptible to misinformation. A bias detection assistant focused on social media-style news could help address this gap.

- **What:** This project develops an NLP tool that scans short-form news content (e.g., from Facebook, Reddit, or TikTok captions) for bias-indicative phrases using sentiment scoring, rule-based lexicons (e.g., weasel words), and lightweight classifiers. The goal is not to fact-check, but to surface patterns in language that suggest rhetorical slant or manipulation.

- **How:** Using pre-trained NLP models (e.g., spaCy or BERT), the system will identify emotionally charged language and vague statements. A scoring system or heatmap will be generated to help users visualise which parts of the text are most problematic. The tool will be tested on synthetic and public datasets, with

evaluation focused on how well it helps users spot manipulation and improves critical reading performance.

**Input Sources**

- **Primary Data:** Social media-style news snippets (e.g., Facebook posts, Reddit threads) collected from public datasets or scraped from verified news sources with associated bias labels.

- **Bias Lexicons:** Hand-crafted and published lists of weasel words, emotionally charged terms, and framing cues used to detect manipulative or slanted language.

- **User Study Data:** Simulated survey responses and comprehension assessments collected from student participants to measure improvement in bias detection.

**Related Work and Lessons Learned**

**1. LIAR Dataset (Wang et al., 2017)**   A labeled dataset of short political statements from PolitiFact.com that includes fine-grained bias and truthfulness labels. It is commonly used to train fake news and bias detection models.

*Key Insight:* The dataset's short-form, headline-style format matches social media structures, but lacks Singapore-specific linguistic features.

**2. Fake News Challenge (FNC-1)**   A challenge dataset focused on stance detection and bias in news headlines. FNC systems identify whether a headline agrees, disagrees, discusses, or is unrelated to a given claim.

*Key Insight:* These models focus on relational understanding between claims and articles — a useful approach for understanding framing, but too heavy for sentence-level bias cues.

**3. Related Research Literature**

- **Baly et al. (2018)** — *"Predicting Factuality and Bias in News Media"*. Joint model for detecting news bias and factuality at the source and article level. Link to paper

- **Horne et al. (2017)** — *"Just What Do You Think You're Doing, Dave?" A Dataset for Detecting Stance and Bias* Explores how lexical and syntactic features signal bias and subjectivity. Link to paper

# 2 Explainable Deep Learning for Breast Cancer Detection in Singaporean Women

**Target Audience**

Healthcare researchers, diagnostic support teams, and public health agencies in Singapore aiming to improve early detection of breast cancer through explainable AI solutions tailored for the local population.

**Template Chosen**

3.2 Project Idea 2: Deep Learning Breast Cancer Detection

**Aim**

To develop an explainable deep learning model that detects early-stage breast cancer from mammography images in Singaporean women with dense breast tissue, providing interpretable visual explanations to support radiologists in difficult screening scenarios.

**Objectives**

1. **Develop a CNN-based model for early-stage breast cancer detection in women with dense breast tissue.** The model will be trained and fine-tuned using mammogram datasets (e.g., DDSM) with a focus on classifying early-stage tumors that are often missed due to high breast density. *Outcome:* Achieve at least **85% sensitivity** and **80% specificity** in identifying early-stage cancer cases in high-density mammograms on a test set.

2. **Adapt the model to simulate Singapore-specific screening scenarios.** Using epidemiological reports and screening guidelines from the Singapore Cancer Society and MOH, data will be filtered and augmented to reflect age (40–55), density profiles, and late-stage first presentations common in the local population. *Outcome:* Demonstrate that the model maintains strong performance (**no more than 10% drop in sensitivity**) when tested on these locally simulated conditions.

3. **Incorporate and evaluate Grad-CAM-based visual explainability to support radiologist interpretation.** Generate heatmaps that highlight regions contributing to classification decisions. These will be validated against tumor annotation masks and assessed for clinical plausibility. *Outcome:* At least **70% alignment** (IoU or pixel-wise overlap) with ground truth regions, and (if possible) qualitative validation from radiologist or test user feedback indicating usefulness in diagnosis.

**Justifications**

- **Why:** Breast cancer is the most common cancer among women in Singapore. Despite nationwide screening programs, some groups—especially those with dense breast tissue or lower awareness—still face diagnostic delays. Radiologists face increasing image loads and diagnostic complexity, particularly in borderline or non-obvious cases. Enhancing diagnostic accuracy and consistency through AI could help address these challenges.

- **What:** This project aims to improve the early detection pipeline by developing a CNN-based model trained to classify high-resolution medical images. The goal is not to replace clinicians, but to support them in identifying patterns in challenging cases. It also serves as a scalable tool that could support screening centres with limited specialist manpower.

- **How:** Using Python-based tools (e.g., TensorFlow or PyTorch), the system will be trained on open datasets like BreaKHis or DDSM. Grad-CAM or similar methods will be used to visualize important image regions that influence predictions, aiding interpretability. Evaluation will focus on sensitivity in early detection and potential alignment with Singapore's national screening goals. Case profiles will reflect local demographic characteristics based on data from the Singapore Cancer Society or MOH reports.

**Input Sources**

- **Image Data:** Public datasets such as BreaKHis (histopathology images) and DDSM (digitized mammograms) will be used to train and evaluate the model.

- **Singaporean Screening Context:** Case profiles and diagnostic guidelines from local sources such as the Singapore Cancer Society and MOH reports will guide the evaluation and simulate realistic diagnostic scenarios.

- **Annotations:** Tumor boundary masks or diagnostic labels provided in the datasets will be used for classification and heatmap validation.

## Related Work and Lessons Learned

**1. Google Health / DeepMind — AI for Mammogram Screening** In 2020, Google Health and DeepMind published a high-impact study on an AI system trained on over 90,000 mammograms across the US and UK. The model demonstrated reduced false positives and false negatives compared to radiologists, and used saliency maps to explain predictions.

*Lesson Learned:* High-quality AI models can match or outperform clinicians in controlled studies, but generalizability to new populations is a key concern. This highlights the importance of regional data adaptation and validation — especially in Singapore's multi-ethnic context.

[Link]

**2. Zebra Medical Vision — Scalable AI Screening Solutions** Zebra Medical Vision has developed FDA-cleared tools for automated cancer detection, including breast cancer, based on medical imaging. Their systems are integrated into national screening programs and offer cloud-based diagnostics.

*Lesson Learned:* Scalable AI solutions for screening rely on not only technical accuracy, but also healthcare infrastructure integration and accessibility. Singapore's screening efforts (e.g., BreastScreen SG) can benefit from similar AI-enabled triage tools.

[Link]

**3. Shen et al. (2019) — Deep Learning on Mammography**  This academic study evaluated convolutional neural networks (CNNs) for classifying mammograms and showed improved sensitivity and specificity in screening tasks.

*Lesson Learned:* CNNs can enhance diagnostic performance, but their reliability depends on high-quality, well-annotated training datasets — which may be limited in local contexts. The use of transfer learning or dataset augmentation is essential.

[Link]

**4. Salim et al. (2024) — *Use of AI for Breast Cancer Screening in Clinical Practice***  This 2024 study published in Nature Medicine evaluated the real-world performance of an AI-supported breast cancer screening system across 58,000 mammograms in Sweden. The AI assisted radiologists by flagging suspicious regions and helped reduce workload without compromising diagnostic accuracy.

*Lesson Learned:* AI can be safely and effectively deployed in national screening programs when integrated into radiologist workflows. Rather than replacing clinicians, this system enhanced efficiency and maintained diagnostic quality — an approach directly aligned with this project's assistive, explainable AI goals for Singapore.

[Link]

## 3. Related Research Literature

- **Shen et al. (2019) — *"Deep Learning to Improve Breast Cancer Detection on Screening Mammography"***  A landmark study using convolutional neural networks (CNNs) to classify screening mammograms, showing improved sensitivity and specificity over radiologists in some scenarios. [Link to paper]

  *Local Relevance:* Demonstrates CNNs' potential in supporting radiologists, especially in large-scale screening contexts. However, replication in Asian or Singaporean datasets is necessary due to population and imaging differences.

- **Raghu et al. (2019) — *"Transfusion: Understanding Transfer Learning for Medical Imaging"***  Investigates the effectiveness of transfer learning for medical image tasks. Highlights limitations in using ImageNet-pretrained models directly for clinical imaging. [Link to paper]

  *Local Relevance:* Encourages the use of domain-specific pretraining or fine-tuning when building CNNs for mammography or histopathology data — an approach our project will follow for Singaporean diagnostic cases.

- **Ching et al. (2018) — *"Opportunities and obstacles for deep learning in biology and medicine"***  A comprehensive review of deep learning use in healthcare, including the importance of interpretability, data bias, and clinical collaboration. [Link to paper]

  *Local Relevance:* Provides critical guidance on designing explainable AI systems that gain clinical trust — reinforcing our use of Grad-CAM and alignment with Singaporean radiological practice.

- **Cheplygina et al. (2019) — *"Not-so-supervised: A survey of semi-supervised, multi-instance, and transfer learning in medical image analysis"***  Discusses strategies for dealing with limited labelled medical data, which is common in local hospital datasets.

*Local Relevance:* Highlights techniques like weak supervision or few-shot learning, which could be applied when Singapore-specific datasets are small or difficult to annotate fully.

# 3 Orchestrated AI System for Early Cognitive Decline Detection in Singaporean Seniors

**Target Audience**

This project is designed for elderly Singaporeans, their caregivers, general practitioners, and community health workers who play a role in cognitive screening and early intervention. Additionally, it targets public health agencies and eldercare services seeking scalable, accessible tools to support dementia prevention initiatives.

**Template Chosen**

3.2 Project Idea 2: Deep Learning Breast Cancer Detection

**Aim**

To develop a multimodal artificial intelligence system that orchestrates voice, language, and behavioural data inputs to screen for early signs of cognitive decline, particularly Alzheimer's-type symptoms. The system will be designed for non-clinical settings and tailored to the speech patterns, language diversity, and lifestyle characteristics of Singapore's ageing population.

**Objectives**

1. Build and evaluate an AI system capable of detecting early cognitive impairment with at least 75% classification accuracy and 80% sensitivity, by processing multimodal inputs (speech, text, lifestyle).

2. Assess system performance using both public datasets and simulated Singapore-style inputs that incorporate multilingual speech and culturally relevant lifestyle markers.

3. Apply explainable AI tools (e.g., SHAP, LIME) to identify the most influential features and ensure transparent, interpretable predictions that support human trust in the screening process.

**Input Sources**

The system will utilize three key types of input data:

- **Speech:** Short audio recordings elicited through structured tasks such as picture description, storytelling, or memory recall. From these, acoustic features such as pitch variability, pause duration, jitter, and articulation rate will be extracted using tools like Librosa or OpenSMILE.

- **Text:** Transcribed speech will be analyzed for linguistic features, including lexical richness, grammatical complexity, coherence, and pronoun usage. Tools like spaCy or BERT-based embeddings will support natural language processing and feature extraction.

- **Behaviour:** Simulated lifestyle indicators will include daily memory lapses, medication adherence, sleep quality, and routine disorientation — modeled from survey data reflecting common early dementia symptoms in Singaporean seniors. These will be encoded as structured categorical or ordinal inputs.

## Justifications

- **Why:** Dementia is one of the most pressing health concerns in Singapore, with the number of affected individuals projected to triple by 2030 due to rapid population ageing. Current diagnostic practices are clinic-based and rely heavily on specialist assessments, which are not scalable for early detection at the community level. Moreover, language diversity and cultural norms in Singapore pose barriers to conventional tools developed in Western contexts. An AI-based system that is non-invasive, multilingual, and explainable could facilitate earlier interventions and broaden screening coverage.

- **What:** This project aims to orchestrate multiple AI subsystems — for speech, text, and behavioural analysis — into a unified model that can identify early cognitive decline. Unlike most existing tools that rely solely on linguistic or medical data, this system leverages multimodal signals to increase robustness. In doing so, it aligns with the broader Smart Nation vision to integrate AI into preventive healthcare services.

- **How:** Pre-trained models like Wav2Vec will be used for extracting acoustic embeddings from speech, while transformer-based language models (e.g., BERT) will analyze transcribed text. Behavioural data will be treated as structured tabular input. All extracted features will be concatenated and passed into a classifier (e.g., Random Forest, XGBoost, or a shallow neural network). Public datasets such as DementiaBank and the ADReSS Challenge will serve as primary sources for training and validation. To simulate local applicability, synthetic data reflecting Singaporean language habits and lifestyles will be crafted. Explainability will be achieved using SHAP to identify and visualize which input modalities and features contributed most to each classification outcome.

## Related Work and Lessons Learned

**1. ADReSS Challenge (2020–2021)**   The Alzheimer's Dementia Recognition through Spontaneous Speech (ADReSS) Challenge provided a standard benchmark for Alzheimer's classification using speech only. Various models were tested on their ability to predict cognitive status using audio and linguistic features without relying on manual annotation.

*Lesson:* The challenge demonstrated that short speech tasks can yield effective biomarkers of dementia. However, most models were trained on native English speakers in Western settings, limiting generalizability. For use in Singapore, the system must handle local English varieties and code-switching patterns common among older adults.

**2. DementiaBank (Pitt Corpus)**   This dataset includes structured tasks like the "Cookie Theft" picture description, along with audio and transcripts of individuals with and without Alzheimer's. It has been widely used for extracting syntactic and semantic markers of decline.

*Lesson:* Structured and semi-structured tasks are repeatable and clinically useful, making them ideal for AI-driven screening. However, task content and vocabulary may need cultural adaptation to avoid misinterpretation by local seniors.

## 3. Related Research

- **Luz et al. (2021)** — *Detecting Cognitive Decline Using Speech Only: The ADReSSo Challenge.* Explores deep learning models using acoustic-only features for dementia detection. Link

- **Fraser et al. (2016)** — *Linguistic Features Identify Alzheimer's Disease in Narrative Speech.* Demonstrates key NLP features that separate Alzheimer's from control speech. Link

- **Haider et al. (2023)** — *Multimodal Detection of Dementia with Text, Audio and Visual Cues.* A multimodal deep learning approach that integrates linguistic, audio, and visual signals. Link