# Kien Pham

ki3n98@gmail.com (714) 592 6510 Westminster, CA Kien Pham Portfolio

## Education

**Bachelor of Science in Computer Science** 08/2024 – 12/2026
*California State University, Long Beach* Long Beach, CA

## Experience

**Thomson Reuters Special Services** 05/2026 – 08/2026
*Data Science Intern* Remote

- Built and benchmarked **86 text-classification models** (11 embedding methods × up to 10 classifiers, plus a zero-shot LLM baseline) to flag suspected juvenile victims in escort ads.
- Fused a text signal (OpenAI text-embedding-Large + logistic regression) with the production image classifier, lifting recall 0.620 → 0.795 and F1 0.50 → 0.64 at unchanged precision.
- Showed that simple linear models generalize across years and quantified **concept drift** as an early-warning metric predicting F1 decline → data-driven retraining trigger.
- Made predictions auditable for law enforcement using **SHAP and TF-IDF + LR coefficients** trained on 2024-only, 2025-only, and pooled data, surfacing durable linguistic signals that held across years.
- **Presented findings company-wide** to non-technical stakeholders, translating model results into a shipping recommendation for text as a second detection signal alongside Spotlight's ~150K ads/day intake.

## Research

**Vision-Based Distracted Driving Detection** 02/2026
*Published & presented at the Sixth Annual CS Conference for CSU Undergraduates*

- Co-authored and presented a benchmark of **Faster R-CNN, SSD, and YOLOv11** for real-time driver-distraction detection on the Roboflow Mobile Detection dataset, selecting YOLOv11 for deployment at 0.604 mAP@0.50:0.95 and 86.1 FPS on a consumer RTX 3080.

## Projects

**Veriface** 06/2026
*2026 CECS Senior Design Award, Outstanding Achievement in Applied AI & ML*

- Built the face-matching pipeline: MTCNN detection → FaceNet (InceptionResNetV1) 512-dim L2-normalized embeddings, cosine best-match against per-event enrollment, served through FastAPI with results in PostgreSQL.
- Implemented JWT/bcrypt auth with a self-designed 5-role RBAC engine (role hierarchy + assignment rules) and audit-logged admin actions — sole author of the permission module.
- Designed the live attendance feed with a custom WebSocket ConnectionManager keyed by session, broadcasting check-in events to active dashboards.
- Shipped 160+ pytest tests (auth, RBAC, events, matching, liveness) and a reproducible Docker/CUDA stack for judges' machines.

## Awards

- **CECS Senior Design Award, Outstanding Achievement in Applied AI & ML** (2026) — VeriFace
- **Best Overall, Marina Hacks 5.0** (2025) — 911 Operator Assistance
- **Best Overall, Beach Hacks 8.0** (2025) — Shape & Sign
- **Best Project, AIS Datathon 2026**

## Skills

- **Machine Learning & AI:** NLP (word/contextual embeddings, text classification), Computer Vision, model fine-tuning, concept drift, logistic regression, Prompt Engineering, LangChain, Multi-agent orchestration, Whisper / faster-whisper
- **Frameworks & Tools:** PyTorch, TensorFlow, Scikit-Learn, FastAPI, Django, Next.js, Docker, Git/GitHub
- **Data:** PostgreSQL, data collection & preprocessing, data labeling, data evaluation, Matplotlib