Prashant Shekhar, PhD

Business Data Science | Causal Inference | Experimentation | Statistical Decision Systems
Daytona Beach, FL | shekharp.erau@gmail.com | +1 (716) 817-4196 | Google Scholar

PhD-trained data scientist and applied mathematician with 10+ years of experience designing statistical decision and machine-learning systems for policy evaluation, experimentation, uncertainty quantification, and applied analytics. Research focuses on causal inference with incomplete evidence: support-aware off-policy evaluation, robust experiment design under interference, privacy-constrained incrementality measurement, and conformal uncertainty for operational pacing decisions. Experienced translating ambiguous business and scientific questions into auditable analytics pipelines, decision rules, technical narratives, and stakeholder-ready recommendations.

Decision Science Agenda

- **Causal and statistical inference**: decision systems for incomplete evidence: off-policy evaluation, replay estimands, robust experiments, incrementality, support diagnostics, guardrails, and launch-readiness.
- **Interpretable machine learning**: structured and transparent models for decision support: sparse learning, multiscale approximation, time-series reconstruction, uncertainty estimates, and model validation.
- **Privacy, AI, and cybersecurity**: decision reliability under privacy, trust, and adversarial constraints: privacy-preserving learning, signal loss, secure distributed learning, robustness, and authentication.
- **Anomaly detection and adoption**: operational decision systems for rare events, defects, domain shifts, and deployment constraints in manufacturing, biomedical monitoring, autonomous systems, and scientific analytics.
- **AI-assisted software development**: local-first research software for lab operations: deterministic tools, read-only infrastructure agents, conversational interfaces, model control, and deployment automation.

Experience

Embry-Riddle Aeronautical University Daytona Beach, FL
Assistant Professor of Data Science / Applied Mathematics, tenure-track *Aug 2021 – Present*

- Lead a research program in statistical decision systems for business and platform analytics, including causal policy evaluation, online experiment design, incrementality measurement, and decision-calibrated uncertainty.
- Develop support-aware offline evaluation pipelines that combine replay estimands, off-policy evaluation, conservative lower bounds, guardrails, subgroup checks, sensitivity analysis, and out-of-time validation.
- Design experimentation frameworks for settings where treatment effects can propagate through shared budgets, producer exposure, inventory, graph spillovers, temporal carryover, or other interference channels.
- Build decision rules that separate promising findings from deployable recommendations, with outputs such as certified policies, rejected alternatives, unresolved candidates, and validation shortlists.
- Led multiple funded AI/data-science research projects totaling $200K+ as Principal Investigator, supervising student teams across statistical modeling, machine learning, cybersecurity, autonomous systems, and scientific analytics.
- Teach and mentor undergraduate and graduate students in data science, machine learning, applied mathematics, statistical reasoning, model evaluation, and technical communication.

Tufts University Medford, MA
Data Scientist *Sept 2019 – Aug 2021*

- Built statistical learning and machine-learning methods for noisy, high-dimensional, and irregularly sampled scientific and biomedical datasets.
- Developed sparse multiscale learning and localized penalized-spline approaches for data reduction, time-series reconstruction, outlier detection, uncertainty estimation, and predictive modeling.
- Collaborated with domain scientists and engineers to turn uncertain measurement problems into reproducible analysis workflows, validated models, and peer-reviewed publications.
- Applied machine learning to rare-event biomedical detection, including label-free flow-cytometry signals and circulating tumor cell cluster detection in whole-blood data.

Agenda Area 1: Causal and Statistical Inference

- ***Decision Support for Marketplace Policies under Incomplete Evidence: From Replay to Launch Readiness***
  arXiv 2026
  - Reframed offline marketplace analysis as launch-readiness triage: choose launch, online validation, rejection, or unresolved rather than chasing the highest offline score.
  - Combined replay estimands, off-policy evaluation, guardrails, support diagnostics, subgroup checks, sensitivity analysis, and out-of-time validation to separate statistical signal from deployable evidence.
  - Found strong offline gains – 47.7% replay lift, 45.8% conservative lower-tail lift, and 43.9% out-of-time lift – but recommended online validation because key causal assumptions were still unresolved.
- ***Support-aware offline policy selection for advertising marketplaces***
  arXiv 2026
  - Turned policy selection into a shortlist decision under support limitations: identify which candidates are credible enough to test online.

- Used simultaneous lower bounds, support gates, regret checks, and subgroup non-harm diagnostics to manage multiple policies without overstating causal certainty.
- Reduced a 19-policy catalog to a two-policy validation shortlist, with 47.66% replay lift, 40.71% simultaneous lower-bound lift, 43.87% out-of-time lift, and non-harm checks across 44 segments.

- ***Choosing Online Experiment Designs under Interference in Ads, Recommendations, and Member-Experience Systems***
  2026
  - Framed experiment design as a decision made before the analyst knows the dominant interference channel, rather than as a default choice between user-level and cluster-level randomization.
  - Linked design choice to business risks such as spillovers, budget contamination, producer exposure, carryover, variance, MDE, implementation cost, and estimand mismatch.
  - Compared six deployable design families and selected different robust designs across settings, with risk scores of 1.295, 2.105, and 2.240 under distinct interference profiles.

- ***Privacy-Robust Incrementality Measurement for Advertising Systems under Signal Loss***
  2026
  - Treated incrementality measurement as a causal-certification problem under privacy-driven signal loss: decide which lift claims remain certified, rejected, or unresolved.
  - Stress-tested match loss, attribution loss, aggregation thresholds, reporting noise, simultaneous uncertainty, and segment granularity so privacy constraints become explicit decision inputs.
  - Evaluated 2.0M advertising rows and a 64K-row randomized marketing experiment; fine-segment certification fell from 100% to 36.0% while suppression rose to 39.5%, clarifying where claims became underpowered.

- ***Decision-Calibrated Conformal Uncertainty for Pacing Decisions in Streaming Advertising***
  arXiv 2026
  - Made uncertainty actionable by calibrating intervals to the downstream pacing decision, not merely to generic forecast residuals.
  - Separated prediction accuracy from decision reliability through guardrails for budget pacing, exposure load, and operational constraint violations.
  - Reduced conformal uncertainty radii from 7236.7 to 18.4 and from 4629.4 to 278.6 in replay studies; reduced held-out violation rate from 16.7% to 3.3% with zero budget and member-load violations in one setting.

Agenda Area 2: Interpretable Machine Learning

- ***A forward–backward greedy approach for sparse multiscale learning***
  CMAME 2022
  - Connected interpretable ML to decision science by producing sparse models that preserve the signal needed for prediction while removing redundant observations.
  - Made model complexity auditable through feature-driven multiscale basis selection, so accuracy, approximation error, and data reduction can be traded off transparently.
  - Demonstrated 62%–80% compression on structured benchmark data and up to 98.6% compression on high-volume sensor data while preserving predictive structure.

- ***Hierarchical Regularization Networks for Sparsification Based Learning on Noisy Datasets***
  arXiv/stat.ML 2019
  - Framed noisy-data modeling as a sparsification decision: keep enough observations to support reliable prediction while removing redundant or unstable detail.
  - Used hierarchical regularization and GCV-style model selection to make the bias–variance–compression tradeoff explicit and statistically inspectable.
  - Selected sparse representations with 0.68 and 0.43 compression in synthetic studies and reached 0.91 compression in the noisiest time-series example while retaining uncertainty intervals.

- ***Hierarchical Data Reduction and Learning***
  arXiv/stat.ML 2019
  - Built interpretable data-reduction models for settings where downstream decisions need fast prediction, uncertainty statements, and clear approximation error.
  - Used multiscale representations with confidence and prediction intervals so compressed models remain statistically accountable rather than purely heuristic.
  - Demonstrated a scale-6 approximation using 23% of observations with 2-norm error below 1e-5, showing that decision-relevant structure can be retained with a much smaller support set.

- ***SVD enabled data augmentation for machine learning based surrogate modeling of non-linear structures***
  Engineering Structures 2021
  - Translated expensive simulation workflows into surrogate modeling decisions: determine which augmented scenarios are informative enough to improve prediction under limited data.
  - Used structured SVD-based augmentation to preserve dominant modes of variation, improving interpretability over purely synthetic black-box data expansion.
  - Showed the first SVD direction captured 91% of variation in a motivating example and evaluated four ML surrogate families on unseen scenarios, supporting faster what-if analysis.

- ***ALPS: A Unified Framework for Modeling Time Series of Land Ice Changes***
  IEEE TGRS 2021
  - Turned irregular, noisy, multi-source time series into decision-grade monitoring outputs: reconstructed trajectories, outlier flags, uncertainty bands, and derivative estimates.
  - Used localized penalized splines and confidence-based screening to make measurement quality and temporal uncertainty visible before acting on trend estimates.
  - Supported 10-day reconstructions, 99% t-confidence outlier screening, and uncertainty estimates at scale for more than 100,000 repeatedly calculated time series.
- ***Glacier surge monitoring from temporally dense elevation time series***
  The Cryosphere 2025
  - Generalized the time-series framework into an operational monitoring problem: detect meaningful regime shifts from long, noisy, irregular measurement records.
  - Balanced smoothing, outlier handling, and temporal resolution so change detection could support review and escalation rather than a single retrospective plot.
  - Produced monthly time series from a 20-year high-noise record covering more than 100 monitored entities, demonstrating scalable decision support for regional change detection.

Agenda Area 3: Privacy, AI, and Cybersecurity

- ***Secure Distributed Learning for CAVs: Defending Against Gradient Leakage with Leveled Homomorphic Encryption***
  arXiv 2025
  - Framed privacy-preserving learning as a deployment decision: how much gradient protection can be added before model utility or operational cost becomes unacceptable.
  - Compared selective and full encryption strategies so privacy risk, architecture choice, communication rounds, and accuracy loss can be evaluated together.
  - Evaluated 60,000-image distributed learning with three clients, 50 communication rounds, and four model families from 3.4M to 21.8M parameters, identifying architectures that remained robust under full encryption.
- ***Advancing Practical Homomorphic Encryption for Federated Learning***
  arXiv 2025
  - Converted homomorphic-encryption configuration into a privacy–utility decision, linking encryption ratio, additive noise, model size, and convergence stability.
  - Connected theoretical guarantees with empirical model selection so privacy controls can be chosen as governed parameters rather than ad hoc safeguards.
  - Identified 1e-6 as a practical additive-noise threshold; final losses at that setting matched or nearly matched the noise-free baseline across models from 88K to 11.2M parameters.
- ***Do Adversarial Patches Generalize? Attack Transferability Study Across Real-Time Segmentation Models in Autonomous Vehicles***
  IEEE SPW 2025
  - Treated adversarial robustness as model-risk assessment: decide whether a model is safe to deploy after stress testing likely failure modes.
  - Separated same-model vulnerability from cross-model transfer risk, which helps teams choose validation depth, model diversity, and monitoring requirements.
  - Tested seven real-time segmentation models; targeted attacks reduced mIoU from 0.8996 to 0.7914, 0.6844 to 0.2991, and 0.8118 to 0.4491 in representative cases.
- ***Radio Frequency Fingerprinting Authentication for IoT Networks Using Siamese Networks***
  2025
  - Recast device authentication as a known/unknown decision problem with explicit acceptance and rejection behavior rather than simple closed-set classification.
  - Used similarity learning to support auditable thresholding, enabling a practical decision rule for admitting known devices and flagging unknown transmitters.
  - Achieved 98.25% overall accuracy on a 6,774-frame test set and correctly detected 3,922 of 3,984 unknown-device samples.
- ***Adversarial-Resilient RF Fingerprinting: A CNN-GAN Framework for Rogue Transmitter Detection***
  2025
  - Extended authentication into adversarial decision-making: distinguish genuine, rogue, and synthetic mimicry attempts under distribution shift.
  - Combined discriminative modeling with synthetic rogue generation so the decision boundary is tested against harder negatives, not only observed devices.
  - Achieved rogue-device F1 of 0.9871, detecting 96.7% of rogue samples and 97.6% of genuine samples in a real-plus-synthetic evaluation.

Agenda Area 4: Anomaly Detection and Adoption

- ***Deep convolutional neural network-based automated optical inspection for aerospace components***
  Digital Engineering 2025

  - Framed visual inspection as an operational decision system for scarce and imbalanced defect labels: classify, review, or escalate quality risk.
  - Combined augmentation, classical ML, deep CNNs, and hybrid models to evaluate whether automation could support repeatable inspection under limited data.
  - Achieved up to 99.68% detection accuracy after expanding an approximately 800-image inspection dataset into 8,000 balanced training images and comparing 18 model variants.
- ***Object Shape Error Modelling and Simulation During Early Design Phase by Morphing Gaussian Random Fields***
  Computer-Aided Design 2023
  - Moved manufacturing quality decisions earlier in the lifecycle by simulating shape-error risk before full production data are available.
  - Modeled spatial correlation, local/global deformation, technological patterns, and statistical tolerance constraints so teams can evaluate likely failure modes under uncertainty.
  - Supported what-if analysis for limited-data design settings, turning geometric variation into quantifiable quality-risk scenarios rather than late-stage inspection surprises.
- ***Deep learning-enabled detection of rare circulating tumor cell clusters in whole blood using label-free flow cytometry***
  Lab on a Chip 2024
  - Treated biomedical rare-event detection as a decision problem under extreme class imbalance: maximize useful alerts while controlling false positives.
  - Used deep learning and transfer-style modeling to convert high-throughput measurement signals into reviewable detection events with purity and sensitivity tradeoffs.
  - Achieved 0.78 false alarms per minute, Pearson correlation 0.943 between detected and expected events, 72% purity, and 35.3% sensitivity for rare clusters starting at two cells.
- ***Label-free flow cytometry of rare circulating tumor cell clusters in whole blood***
  Scientific Reports 2022
  - Built a monitoring decision workflow from noisy measurement signals, reducing dependence on invasive or label-heavy procedures.
  - Combined signal features with ML classification so detection thresholds could be evaluated through accuracy, purity, and missed-event risk.
  - Detected approximately 93% of rare clusters larger than two cells with greater than 82% purity and greater than 95% overall accuracy.
- ***PILOT: A Data-Free Continual Learning Approach for Real-Time Semantic Segmentation via Boundary Guidance***
  arXiv 2026
  - Framed continual learning as an adoption decision: add new categories while preserving trusted behavior on existing categories.
  - Used boundary-guided, data-free adaptation to manage the stability–plasticity tradeoff without requiring full retraining or stored historical data.
  - After four sequential class additions, retained 95.6% of base performance, maintained 71.87% overall mIoU, and outperformed fine-tuning by 11.97 percentage points.

Agenda Area 5: AI-Assisted Software Development

- ***AI_lab_manager: Open-source LLM-powered multi-agent lab operations system***
  Research software / GitHub project
  - Designed and deployed multi-agent LLM system for managing distributed infrastructure using control-plane + server agents via natural language interfaces.
  - Worked on prompt optimization and model behavior improvement through iterative evaluation and feedback.
  - Implemented pipelines with iterative feedback, ranking optimization, and output prioritization.
  - Improved efficiency and reliability of system-level information retrieval, personalization, and monitoring.
  - Built reusable LLM pipelines and orchestration systems supporting modular workflows.

Education

The State University of New York at Buffalo — Buffalo, NY
PhD in Computational Data Science; MS in Mechanical Engineering — *Aug 2014 – Aug 2019*

Skills

- **Causal and statistical inference**: causal estimands, randomized experiments, off-policy evaluation, support diagnostics, sensitivity analysis, conformal uncertainty, guardrails, uncertainty quantification, and launch-readiness decisions.
- **Business and product analytics**: experiment design, incrementality and lift measurement, marketplace policy evaluation, subgroup analysis, product metrics, guardrail metrics, dashboard-driven insights, and stakeholder-facing recommendations.
- **Supervised and unsupervised learning**: classification, regression, tree and ensemble models, neural networks, clustering, anomaly detection, dimensionality reduction, representation learning, validation, and error analysis.
- **Data visualization**: Tableau, Matplotlib, seaborn, Plotly, dashboard design, exploratory analysis, metric storytelling, model diagnostics, and decision-facing visual summaries.
- **Causal inference toolkit**: DoWhy, EconML, causal-learn, DoubleML, treatment-effect estimation, DAG-based reasoning, uplift/incrementality workflows, and tutorials developed for applied learners.

- **Coding agents**: Codex, Claude Code, AI-assisted software development, agentic coding workflows, tool orchestration, repository navigation, debugging, documentation, and review.
- **AI models**: LLMs, vision-language models, GenAI, prompt optimization, model behavior evaluation, retrieval/tool-use patterns, and research portfolio visible through Google Scholar.
- **AI stack**: Hugging Face, Ollama, vLLM, local model serving, model switching, LLM lab code, open-model workflows, and inference experimentation.
- **Data science toolkit**: Python, SQL, scikit-learn, statsmodels, PyTorch, NumPy, pandas, SciPy, Jupyter, reproducible analysis, and publication-quality technical workflows.
- **Leadership and communication**: PI experience, student/team supervision, research proposal writing, peer-reviewed publication, technical review, teaching, mentoring, and cross-disciplinary collaboration.