Kevin Wang Ph.D.

Data Scientist and Statistician

US: (+1) 619-215-3725
AU: (+61) 404-955-255
kevinwangstats@gmail.com
Portfolio: kevinwangstats.com

Experience

Now
2021 Nov — **Senior Data Scientist**, *Illumina*

- Led bioinformatics R&D for multiple pioneering cancer companion diagnostics programs. Oversaw high-value data deliverables which directly contributed to multimillion dollar contracts and regulatory milestones.
- Designed, deployed and maintained organizational-wide data science tools and data managment dashboards in Python and R. These tools were foundational infrastructure for high-throughput clinical genomics product development. These tools fundamentally improved the overall efficiency of data reporting within the R&D organization and reduced the data verification processes from several months to hours.
- Facilitated cross-functional collaboration among R&D scientists, software developers, regulatory authorities and pharmaceutical partners to achieve regulatory milestones for IVD and IUO products in both US and EU. This leadership ensured alignment across diverse teams, facilitating successful product approvals.
- Developed machine learning algorithms to analyze large-scale genomics datasets, uncovering clinically-actionable insights that significantly enhanced the performance of oncology and prenatal testing products.
- Spearheaded a comprehensive review of organizational data management practices for the clinical and medical operations teams. Developed and formalized a scalable framework to enhance data velocity and data standardization. This streamlined decision-making and data verification bottlenecks that ultimately fast-tracked critical data deliverables.
- Led and authored a high-impact publication for a clinical oncology product's benefits for blood cancer.

2021 Oct
2020 Jul — **Statistician**, *CSL Behring*

- Direct contributions to the R&D and manufacturing of high-value pharmaceutical products like the COVID-19 vaccines and treatments during the pandemic.
- Designed and formalized statistical process control techniques to standardize practices in the manufacturing and quality control laboratory. These tools were deployed through interactive dashboards which had a wide organizational impact.
- Provided actionable insights to multiple pharmaceuticla products which mitigated significant regulatory risks and achieved significant financial savings, with some up to multimillion dollars production batches.

2020 Jun
2019 Sep — **Research Associate**, *University of Sydney*

- Developed an interpretable machine learning framework for predicting patient clinical outcomes using omics data. Melanoma was a key focus of this research and the key biomarkers were validated and implemented on a cost-effective assay.
- R software packages associated with the research were developed and maintained in a reproducible environment.
- Visiting scholar at Cornell University for 2 months. Orchestrated operations, designed contents and hosted a workshop in open-source single-cell analytics leveraging cloud computing.
- Published peer-reviewed articles in top-tier journals, including PNAS, Nature Digital Medicine and Nature Computational Science.

2019 Mar
2016 Mar — **Postgraduate Teaching Fellow**, *University of Sydney*

- Delivered lectures (200+ students) and tutorials, covering 15 different courses, including statistics, mathematics and data science. Achieved 85% satisfactory rating for end-of-course surveys.
- Authored and designed course content at both the undergraduate and postgraduate levels.
- Mentored and trained junior statistics tutors.

Education

2016–2020 **Doctor of Philosophy in Statistical Bioinformatics**, *University of Sydney*
- ○ Developed machine learning methods to solve statistical challenges in patient clinical outcome predictions and thereby enabling personalised medicine.
- ○ Real-world biomedical data from collaborators were used to motivate and articulate important challenges in current statistical research. The value of this thesis is the synthesis of statistics, clinical implementation and practical deployment.

2012-2015 **Bachelor of Science (Adv. Mathematics) (Hon. I)**, *University of Sydney*
- ○ Major in statistics and financial mathematics.
- ○ The Honours thesis examined functional MRI data and the inferred brain connectivity networks through the use of different statistical association measures.

Skills

python Developed computational modules for regulated products at Illumina to support clinical studies. Highly proficient in writing scripts and notebooks to solve analytical problems for stakeholders. Experience in using and applying `pandas`, `numpy`, `scikit-learn`, `keras`, `pytorch`, and `openai` for a range of projects.

R Highly proficient in scripting and developing packages. Strong ability to use `tidyverse` and `shiny` for the data ETL process to produce visualisations and presentable insights. Author and maintainer of several open-source `R` packages, including `APES`, `learningtower`, `mcvis` and `scMerge`.

git & CI/CD & docker Highly proficient in a reproducible environment that supports business operations and software development requirements. Experience in holding reproducible workshops with excellent participant feedback.

Applied statistics Highly experienced at the research level and consulting level. Ability to transform business and cross-disciplinary problems into a statistical framework and deliver solutions. Strong experience with model selection to extract explainable insights from data.

Statistical modelling Highly experienced at the research level and consulting level. Specialisation in predictive modelling for high-dimensional data.

Other skills SQL, cloud and dashboard development for business use.

Volunteering

2021-Present **Mentor**, *Monash University*
Supervised and mentored three Master of Data Science students with the students successfully completed their research projects. Developed high-quality R-CRAN package and published academic journal article.

2021-2023 **Councillor**, *Statistical Society of Australia*
Served as the Assistant Secretary and Communication Officer in the NSW and Victorian Branch of the Society. Organised public events, designed automated management systems, responsible for member communications and social media account maintenance.