

Curriculum Vitae

Personal Information

Name: Corrin Lakeland
Address: 21 Harbourview Road, Lower Hutt
Email: lakeland@acm.org
Web: <http://go.org.nz/~corrin/>
Date of Birth: 19/3/1976
Telephone: home: (04) 586 9343
mobile: (021) 467 784

Professional Expertise

I have an extensive and wide ranging background in the field of Computer Science and Statistics, allowing me to solve a wide variety of technical problems. I love providing innovative solutions for people that work in the real world. I work well both managing a problem from start to finish by myself as well as coordinating a technical team.

Academic Background

Bachelor of Science (Hons), University of Otago, 1997
PhD, University of Otago, 2006

Employment History

2006-present : Senior Analyst (Datamine)
2005 (May to August): Assistant Research Fellow
2001-2005 (April): Completing PhD, a new algorithm for backoff
2000: Sole tutor for 3rd year Artificial Intelligence course, Victoria University
1999: Lecturer at EIT in Data Communications
Prior to 1999: Numerous tutoring positions at Otago University

Professional Experience

My latest projects involve developing innovative customised data solutions for businesses including major multinational blue chip companies. Three case studies are given below to give a sample of the larger jobs I have completed.

Case study 1: Address Tools

A data consulting company frequently had to deal with addresses and had developed a number of tools to assist with matching, postcoding, geocoding, verification and the like. As the volume of work increased they found their tools were inadequate, leading to mistakes such as delivering mail

intended for Queen Street, Kerikeri to Queen Street, Auckland and Manukau City to the Manakau farming community near Kapiti.

I analysed the different situations in which the tools were used and developed end-to-end replacements that are faster, substantially more accurate and able to flag where they may have made a mistake for manual attention.

These tools outperform all others on the market and rapidly led to the company going from processing addresses as a peripheral operation to the market leader. The tools were unusual in that they were being used by technical experts and so my design concentrated on flexibility over ease-of-use, providing features such as detecting duplicates, merging different address databases, formatting addresses, printing reports of common errors, and many others with new functions easily added. The total savings for their clients already exceeds three million dollars in just six months of use.

Case study 2: Customer Retention

A large supermarket knew many of its customers were regulars but did not know why they chose to shop there, if they also shopped elsewhere, or anything much else about them. This lack of customer understanding made it impossible for it to provide the services its customers were looking for.

By analysing membership of the supermarket's loyalty programme, I found their customers are generally affluent (relatively insensitive to price changes) and they chose to shop there because it is the closest supermarket to where they live. Just this information was valuable, leading to the supermarket concentrating on added services to attract customers over price specials, but the bulk of the analytic work was modeling the behaviour of all regular customers to detect trends.

The behaviour analysis work is best explained by way of example: one of the supermarket's higher value customers began exhibiting irregular behaviour, spending less most weeks and sometimes not coming in at all. My software noticed this change in behaviour and wrote to the customer, giving them a voucher and asking if the supermarket could do anything more from them. The customer responded that he had started shopping around since his favourite brand of snacks were often out of stock. In response, I modified the supermarket's logistics setup to predict stock levels more accurately and as a result was able to reduce the frequency and duration of product unavailability without substantially increasing warehouse requirements.

Case study 3: Media Effectiveness

A large investment company had a problem. Despite spending several million dollars every year on advertising, it was slowly losing customers. The company felt if it was able to boost calls to its call centre then it would be able to convert them into sales but wanted to do so more efficiently than simply spending more on advertising.

I analysed people contacting the call centre and (age, gender, where they lived) and by looking at the 0800 number called and when they called determined which advertisement they were responding to.

Based on this I found that while prime-time advertisements were the most effective, they were nowhere near effective enough to justify their premium and the company was much better off with targeted specialist advertisements and sponsorship. This has resulted in saving the company quarter of a million dollars a year for the same response, and by continuing with its previous budget, the company has reversed the trend and is now gaining customers. Additionally, I found the potential

customers did not identify with the advertisement campaign's main character and so a new campaign was developed, with the customer model that was built from call centre respondents assisting the new campaign's creative design.

Other examples of work

- Updating many customer databases to the new postcode scheme
- Cleaning and formatting all aspects of an address (leading to higher delivery rates)
- Datamart maintenance for large financial companies, supporting mergers, deduping, data mining, etc.
- Merging datamarts with different designs and formats after a large corporate merger
- Market share and share of wallet calculations
- Developing New Zealand's most accurate SendRight software (address accuracy)
- Developed a tool for finding duplicate entries in a customer database even when every field contains errors. Using this greatly increases the accuracy of customer databases and leads to better customer understanding.
- Developing by far New Zealand's best tool for matching customer addresses against New Zealand Post's 'new movers' list (HPAD), allowing GNAs and other lost customers to be followed.
- Developed a meshblocking tool that allows rapid and accurate profiling of customer bases against the census.
- GIS profiling of companies showing travel distance, customer penetration, etc.
- Designed an advertising offer for a major NZ bank leading to a very successful Christmas, reversing previous trends.
- Designed a donor profile tool for a large NZ charity, allowing them to contact people most likely to donate and so turning a very negative ROI into a profit.
- Developed a model for one of NZ's main telecommunication providers that analysed telephone usage and used it to predict when a customer would churn, allowing the company to contact them in advance and offer enticements to keep them on. Actual increased sales are unknown currently but were provisionally estimated at \$30k/year.
- Developed a mortgagee profile tool for a smaller NZ lender allowing them to personally contact ideal customers and leading directly to a spike in sales.
- Worked with sensitive data, showing maps where claimants are most likely to live, where services are duplicated, and managed service provider expectations during a campaign to contact claimants

Earlier employment history:

Working as a research fellow, I was adding features to a large lisp program for machine translation. This involved becoming familiar with the system in a few weeks. Specific features were building a database of parses and using it to build a probabilistic grammar model, and implementing a test-suite.

Doctoral Thesis:

My doctoral dissertation was in artificial intelligence, with most of the research in unsupervised data mining. A short abstract follows:

I propose a method for optimising the trade-off between informative and learnable constructions in statistical parsing. I implement a grammar which works at a level of

granularity in between single words and parts-of-speech, by grouping words together using unsupervised clustering based on bigram statistics. I begin by implementing a statistical parser to serve as the basis for our experiments. The parser, based on that of Michael Collins, contains a number of new features of general interest. I then implement a model of word clustering, which is the first to deliver vector-based word representations for an arbitrarily large lexicon. Finally, I describe a series of experiments in which the statistical parser is trained using categories based on these word representations.

At Victoria university I was involved in designing and presenting all tutorials in artificial intelligence, as well as marking all assignments. As it was the lecturer's first year working, I also assisted in designing the assignments and setting the level. This is similar to my previous work at Otago.

My work at EIT was working as a tutor. I completely ran a course in data communications and networking (06.523) from developing lecture notes, CAL software, tutorials, through to presenting the material, and setting and marking all assessments. The course covered theoretical aspects as well as including an extensive hands-on component. My goal was to prepare students for designing and managing a computer network for any small to medium business. I also gave several guest lectures in setting up Linux and Apache.

During my time at Otago I have done a lot of tutoring. For instance, I taught COSC231 which teaches intermediate programming. This included explaining basic concepts and debugging techniques. I also taught COSC232 which teaches the scientific process, writing skills, and research methodology, and marked most of the COSC232 student assignments. In the following year I continued to tutor the second year programming course, and was the main tutor for one of the data structures course, marking all the assignments. I also assisted in the design and running of tutorials for the third year Artificial Intelligence paper. Finally, I worked as a senior tutor for the third year programming course which was being offered for the first time. This work included setting and marking assignments.

I have kept myself busy in between and during the jobs discussed above including some programming contracts: designing a simple grammar tool for a publisher, implement an educational game, built simple online-ordering systems, built a web authentication system for accessing restricted information, as well as other computer-related jobs: help set up and manage an ISP, help desk support, building computers, corporate requirements planning, etc. These were minor short-term positions and not worth elaborating on.

Programming and Technical Skills:

I first learned programming over twenty years ago and have been programming extensively for thirteen years. During that time I have completed projects in about a dozen programming languages. In approximate order of familiarity, these are C, C++, SQL, Python, Perl, Lisp, R, Java, SAS, Haskell, Bash, Pascal, Basic, Assembly, and Prolog. I am also familiar with a number of other languages and learn new languages easily. For instance, I was brought in to help fix a broken website that was written in PHP and became comfortable enough in the language to fix the problem within an hour of first using it.

I have worked on extremely small projects where the proposal is given just a couple hours before the job is completed through to relatively large projects. My last thesis involved writing approximately twenty thousand lines of code, and the previous one was a little under ten thousand lines. When working on KDE, I was modifying a system containing over thirty million lines.

I have also tutored several courses involving programming at university: First, second and third year applied programming, second and third year algorithms and data structures, third year artificial intelligence.

During my time using computers I have become familiar with a very wide variety of operating systems and applications. I am familiar with all major applications and operating systems currently used, and learn new applications quickly and easily.

Open Source:

I have also been involved in a number of projects within the Linux community, such as instigating the Maori translation project for Linux, and managing the Maori translation in KDE for a number of years. I have continued to contribute to the development of Debian, one of the largest Linux distributions, principally in the area of spam detection. I have also spent some time working in linux technical support, helping with the problems that the tier one support could not solve.

Refereed Publications:

- 2006 **Lakeland C** Improved Statistical parsing, PhD thesis, University of Otago.
- 2004 **Lakeland C** and Knott A, Implementing a lexicalised statistical parser, Australasian Natural Language Processing Conference, Sydney.
- 2003 **Lakeland C**, Some experiments in computational linguistics, New Zealand Linguistic Society.
- 2001 **Lakeland C** and Knott A, POS tagging in statistical parsing, Australasian Natural Language Processing Conference, Sydney.
- 2001 **Lakeland C**, Statistical Parsing, New Zealand Linguistic Society.

Invited seminars:

- 2005 Neural Networks for Lexical Backoff, Macquarie University (Sydney)
- 2005 Designing a new lexical representation for backoff in parsing, Melbourne University Artificial Intelligence Group
- 2005 Neural Networks for Backoff, Computer and Information Science Departments, University of Otago
- 2000 Statistical Parsing, Computer Science Department, Victoria University
- 2000 Introduction to Computational Linguistics, Linguistics Department Seminar, Victoria University.

Professional Activities and Memberships:

Member of the Association for Computational Machinery (ACM)
Guest lectures for fourth year AI
Peer review of a journal paper (2004)
Programme committee of ALTW (2003)
Programme committee of ALTW (2004)
Programme committee of linux.conf.au (2005)

Interests:

I enjoy playing Go, which is an Asian tactical board game.

I also enjoy cooking for friends, tinkering with computers, helping with Linux projects, and playing other games like chess and othello.