Thomas Debeauvais – 949-836-2761 – tdebeauv@uci.edu – http://www.ics.uci.edu/~tdebeauv/

Data scientist and quantitative user researcher with strong programming skills and intellectual curiosity. Favor simple communicable graphs and domain expertise over complicated black-box models.

Education

2010-2016	PhD in Informatics, University of California, Irvine Relevant courses: Quantitative Methods, Machine Learning, Information Retrieval
2007-2010	MS in Computer Science, ENSEIRB, France
2005-2007	Intensive maths and physics prep school, Amiens, France

Skills

Data & Viz	R, Excel, regressions, decision trees, Weka, SPSS, Tableau (basics)
Programming	Python, SQL, JavaScript (jQuery, Bootstrap), Java (Weka, Hadoop), async. networking, UML

Experience

2010-2016: Grad student researcher at UC Irvine, CA
Mined telemetry data for the game Jelly Splash using SQL, Python, and R. Used generalized linear models to explain correlations between churn, revenues, and domain-specific features such as level difficulty, level hopelessness, and gates asking players to invite a friend. Recommended tweaking or removing the gates.

Analyzed a survey of 2800 World of Warcraft players in Excel to show that a quarter of the player base purposefully maintains a $13/month subscription going although they stopped playing.

Increased AUC of models predicting Wikipedia page vandalism from .93 to .97 by coding and tuning SMOTE over-sampling in Weka random forests. Implemented and ran parameter tuning through Hadoop on Amazon EC2. Showed overfitting appears despite cross-validation when exploring parameter space too aggressively.

2013: Intern at Microsoft Research, Redmond, WA
Aggregated 18 months of race data from millions of Forza Motorsport 4 players. Predicted player skill with linear regressions. Provided feedback to the designers on car usage, progression, retention, and assist configuration. Some of the feedback on assist configuration made it into Forza Motorsport 5.

2012: Intern at PARC, Palo Alto, CA
Aggregated and pulled 7 months of daily data from 1,800 World of Warcraft players in MySQL. Used SPSS for visualization and logistic regressions predicting retention. Recommended releasing new content every 6 months.

2011: Intern at Siemens Corporate Research, Princeton, NJ
Designed and implemented a medical web platform prototype using Spring and jQuery. Designed a JSON REST-like API to expose mashups from several third-party heterogeneous data stores and health records to iOS and web front-ends. A dev team later used the prototype as a starting point.

2010-2014: Lecturer and Teaching Assistant at UC Irvine, CA
Designed and taught a class on software architecture. The class used Python and a small networking library that I built on top of asyncore. Mentored dozens of undergrad projects with respect to requirements gathering (customer interviews), usability (Balsamiq mock-ups, cognitive walkthroughs, think-alouds), software architecture (UML), and code reviews (Python, Java, Objective C, PHP, SQL).

Selected publications

- Gate me if you can: The impact of gating on retention and revenues in Jelly Splash, FDG 2015
- Off with their assists: An empirical study of driving skill in Forza Motorsport 4, FDG 2014
- Retention and progression: Seven months in World of Warcraft, FDG 2014
- Distributed tuning of machine learning algorithms using Map-Reduce clusters, LDMTA 2011