

# LORI A. NEWHOUSE

Houston, TX

[LoriNewhouse100@gmail.com](mailto:LoriNewhouse100@gmail.com)

[281.684.3027](https://www.linkedin.com/in/lorinewhouse)

[www.linkedin.com/in/lorinewhouse](https://www.linkedin.com/in/lorinewhouse)

## PRODUCT DEVELOPMENT ENGINEER FOR PREDICTIVE MODELING

**Business Driven Functionality | Machine Learning | Physics Based | Data Science | Value Generation | Flexible Development**

Work with business leaders, end users, and other stakeholders to identify and articulate drivers for and benefits of computational functionality. Formulate product requirement specifications and implementation plans. Maintain close contact with customers in order to adapt these as business needs evolve during product life cycle. Design, implement, test, and validate functionality. Provide user support and training for roll out. Deliver high quality products in timely and cost-efficient fashion. Core competencies:

**Business Requirements | Customer Focused | Multidisciplinary Team | Product Life Cycle**

**Analytics | C++ | Python | scikit-learn | pandas | OOP | Scrum | Agile**

## PROFESSIONAL EXPERIENCE

**SPRINGBOARD**, Houston, TX

2020 to present

### **Machine Learning Engineer (fellow)**

Pursuing career track certification with planned completion March 2021. Program includes 500+ hours of hands-on course work ([GitHub](#) jupyter notebooks) with weekly 1:1 industry mentor oversight. Topics include statistics, data science, ML and AI algorithms, and cloud technologies. Final capstone project is end-to-end: data wrangling, model development, and deployment.

#### **Data Topics and Tools**

- API, RSS, csv, json, text, time series
- EDA, clean, transform, visualize
- numerical features, scaling
- categorical: ordinal, non-ordinal, high cardinality
- encoding: one-hot, frequency, clustering
- train-test split, over and under sample
- numpy, pandas, regex, matplotlib, seaborn
- SQL, PySpark, Databricks

#### **Modeling Topics and Tools**

- cross validation, performance metrics
- feature selection and importance
- regression, classification, clustering
- anomaly, recommendation, time series analysis
- decision trees, neural networks
- sklearn, scipy, pyod, statsmodels, pmdarima
- xgboost, catboost, lightgbm, graphviz
- imblearn, tensorflow, keras, pycaret

#### **Mini-projects ([GitHub](#))**

- |                              |                                     |                          |                          |
|------------------------------|-------------------------------------|--------------------------|--------------------------|
| • <i>data preparation</i>    | quandl.com api                      | • <i>clustering</i>      | customer segmentation    |
| • <i>data processing</i>     | pandas and json data wrangling      | • <i>trees, boosting</i> | loan worthiness          |
| • <i>data processing</i>     | SQL with Spark (Databricks)         | • <i>anomaly</i>         | retail sales             |
| • <i>data transformation</i> | server logs with Spark (Databricks) | • <i>recommendation</i>  | movies                   |
| • <i>regression</i>          | housing price                       | • <i>time series</i>     | stock market forecasting |
| • <i>classification</i>      | gender                              | • <i>Spark ML</i>        | income category          |

#### **Capstone Project ([GitHub](#))**

- Multi-class (7) classification in environmental science with significant imbalance; 580,000 observations.
- Features: 10 numerical, 2 non-ordinal categorical (1 with cardinality of 42).
- Preliminary modeling work with sklearn (11 models); subsequent work with pycaret (5 models).
- Investigated 5 encoding techniques for high cardinality feature; domain informed reduction to 4 clusters performed best.
- Investigated oversampling with imblearn; increasing 5 small target classes by factor of 3 improved prediction performance.
- Predictions served via python script with error checking of input data and process logging.

**HALLIBURTON**, Houston, TX

2007 to 2020

### **Technical Advisor - Engineering and Petrophysical Applications**

Performed technology development for commercial software to design and monitor coiled tubing oil/gas well intervention jobs. Worked with business leaders to understand drivers and benefits. Prepared requirement, specification, design, testing, and user documents. Coordinated with customers during entire development cycle. Updated activities as business needs evolved.

### **Data Scientist Certification 2019**

- Completed inaugural corporate training program (university based online courses, 30 graduates.)

- Final project predicted drill bit refurbishment category from design specifications and field usage data.
- Proposed project to determine relationships between equipment wear and well intervention job characteristics was 1 of 3 selected for proof-of-concept funding by corporate technology.

#### **Computational Modeling**

- Developed detailed, step-by-step procedure to automatically adjust engineering model parameters to obtain match between predicted and measured values.
- Successfully applied procedure to 5 jobs. Identified characteristics of a 6th job which caused procedure to fail.
- Submitted 2 US patent applications for use of real time calculations to manage and control job execution.

#### **Sensor Data Analysis**

- Analyzed time series data from 8 new job site equipment sensors for 6 jobs each 8-12 hours in length.
- Identified and resolved data synchronization and labeling issues.
- Worked with operations leadership and subject matter experts to understand engineering best practices for adjusting equipment parameters during job execution.
- Identified time periods when operation was not consistent with those practices causing increased equipment wear and increased risk of a safety incident.

#### **Software Development**

- Used standard source control procedures (check out and in, check-in builds, nightly builds, release builds, branching, merging.)
- Followed scrum methodology (daily stand-ups, burn down, sprint planning, sprint reviews, code reviews, testing.)

#### **OBJECT RESERVOIR CORPORATION, Austin, TX**

2002 to 2006

##### **Petroleum Physics Engineer**

Developed and supported software for finite element modeling of petroleum reservoirs in venture capital funded company.

- Determined requirement and implementation details of functionality for opportunities identified by company leadership.
- Planned project scopes to fit customer needs and executive management specified delivery dates.
- Designed and implemented calculations for reservoir initialization, fluid property correlations, and grid adaptivity.

#### **MISSION RESEARCH CORPORATION, Nashua, NH**

1999 to 2002

##### **Scientific and Engineering Technical Staff Member**

Designed, implemented, and tested object-oriented framework to calculate atmospheric effects following nuclear burst.

#### **EXXON PRODUCTION RESEARCH COMPANY, Houston, TX**

1991 to 1999

##### **Senior Research Engineer**

Developed and supported reservoir simulation capabilities (phase behavior, fluid flow, thermal, rock fracture.) Activities included writing new programs, debugging and enhancing existing programs, troubleshooting problems, and composing user guides.

### **DATA SCIENCE TRAINING**

#### **coursera online classes**

**Aug to Oct 2019** (certificates available)

- Introduction to Data Science in Python
- Machine Learning Specialization
- Machine Learning Foundations: Case Study Approach
- Machine Learning: Regression
- Machine Learning: Classification
- Machine Learning: Clustering and Retrieval

#### **other online classes (coursera, edx)**

**2020**

- Applied Plotting, Charting, and Data Representation in Python
- Applied Machine Learning in Python
- Data Analysis with Python
- Computing in Python III: Data Structures

### **EDUCATION**

- **Doctor of Philosophy (PhD), Master of Science (MS)**, Chemical Engineering, University of California, San Diego, CA
- **Bachelor of Science (BS)**, Chemical Engineering, University of Michigan, Ann Arbor, MI