# Transforming Life Insurance with Big Data and Machine Learning

## Introduction

Data and analytics are at the center of the insurance market. When you bundle data and analytics capabilities with capital, you have an insurance company. When you bundle them with demand, you have a broker or an advisor.



According to Nucleus Research, analytics return $13 for every dollar invested. Why, then, is only 12 percent of data leveraged for analytics? (Forrester)

The answer may be that there are several challenges that have been delaying projects related to analytics.

With today's advances in technology, many businesses are taking on transformative projects because the rewards are bigger than the pain.

In this industry brief, we are going to look at how machine learning and big data can help your business overcome the challenges and realize the value trapped in your data.

We are going to address specifically the life insurance business, where adoption of these technologies has been slower than in other sectors.

## What Is Machine Learning

The simplest definition of machine learning is *software that improves its own performance over time when measured against a specified task.*

Machine learning has a broad definition and can mean many different things to many people. However, in its most basic form, it is using past data to solve future problems. What makes machine learning such a popular field is that machine learning models can develop nuanced relationships between variables—beyond what a normal person could find. And, unlike with human skill, pattern detection becomes more pronounced and precise as the data quantity grows. Machine learning is a part of predictive analytics, and it is made up of deep learning and statistical/other machine learning. For deep learning, algorithms are applied that allow for multiple layers of learning more and more complex representations of data. For statistical/other machine learning, statistical algorithms and algorithms based on other techniques are applied to help machines estimate functions from learned examples.

## Predictive Modeling Is Instrumental for Profitable Growth of a Life Insurance Company

If we look at the life cycle of a life insurance policy, we can identify five phases:

- Marketing and risk selection
- Quoting, underwriting, and pricing
- Operations and policy management (billing, payments, notifications, claims management)
- Fraud detection and prevention
- Retention

By making data-driven decisions in each of these phases, the life insurer can maximize revenue, while keeping risk and cost in check.

For example: A marketing campaign can be targeted to a group of prospects who are more likely to respond positively. These prospects are selected from millions of consumers, based on aggregated data about their lifestyle, credit history, prescription-drug databases, and other sources. The price of the policy is also determined by using big data and predictive models that supplement the existing actuarial models. Customer acquisition time and costs go down significantly when you can automate much of the process. Customer satisfaction goes up when they can get a quote and purchase a policy online in 20 minutes, instead of one month.

## Challenges with Current Data Management Systems

With growing data volume and variety of data sources available for analysis, many life insurers struggle to figure out the best way to process, gain visibility into, and derive value from more data. Most rely on data sampling, which reduces the accuracy of their risk models and prohibits exploration. Their data is stored in silos, by region or by business unit, on proprietary architectures that are costly to expand. Data correlation is challenging; there are many inconsistencies and it's very difficult, if not impossible, to add new, unstructured data sources.

These architectures have been built for batch processes and cannot be adapted efficiently to the speed of real-time online customer interactions.

## Unified Data Management

The solution we are proposing is to create a unified data platform, with capabilities of unlimited scalability, real-time ingest, high performance processing, and solid security.

On top of this platform, life insurers can then add advanced analytics, such as machine learning. The platform, based on Cloudera's software and its partners' products, also offers self-service and data management tools for actuaries and other analysts, which enables data exploration and cuts the time it takes to get results.

By combining related or even disparate data sets, actuaries and underwriters can reveal patterns, correlations, or causal relationships that, when translated into opportunity or risk, can provide life insurers with a valuable competitive advantage.

With Cloudera, an insurer can now figure out and solve problems as they happen, or even prevent them before they happen. Furthermore, advanced analytics tools deployed as part of the enterprise data hub also provide the insurer's agents with customized recommendations for clients to offer policies based on information gathered in real time on current prospects—essentially monetizing Hadoop's capabilities delivered and supported by Cloudera Enterprise Data Hub Edition.

## Machine Learning as a Competitive Advantage

Life insurance companies typically do not underwrite customers who have suffered and survived from serious diseases like cancer. Doing so would require a long and expensive medical assessment process. With DataRobot, a life reinsurer can use medical history and conditions to accurately predict the risk of underwriting a serious disease survivor. The insurer can identify which customers have good health prospects and directly underwrite them without further medical assessment, leading to more customers and reduced medical costs. Accurate predictive models, enabled by machine learning, have direct impact on the top line and the bottom line of an insurer. With DataRobot, the models are validated and can stand scrutiny from regulators and auditors. Models can be easily deployed across different products, regions, and functions, creating an agile platform that can adapt to business requirements quickly. Another very common predictive analytics use case in underwriting is triage based on lifestyle-related risk factors (e.g., smoking status).

## About DataRobot

DataRobot captures the knowledge, experience, and best practices of the world's leading data scientists, delivering unmatched levels of automation and ease of use for machine learning initiatives. DataRobot enables users to build and deploy highly accurate machine learning models in a fraction of the time by automating the entire modeling lifecycle.

## About Cloudera

Cloudera delivers the modern platform for data management and analytics. The world's leading organizations trust Cloudera to help solve their most challenging business problems with Cloudera Enterprise, the fastest, easiest, and most secure data platform built on Apache Hadoop. Our customers can efficiently capture, store, process, and analyze vast amounts of data, empowering them to use advanced analytics to drive business decisions quickly, flexibly, and at lower cost than has been possible before. To ensure our customers are successful, we offer comprehensive support, training, and professional services. Learn more at cloudera.com.

Cloudera_IndustryBrief_Tranforming_Insurance_102