Case Study for Problem Definition and Target Setting

Source: Adapted from Doing Data Science

Case Study: Real Direct

Doug Perlson, the CEO of RealDirect, has a background in real estate law, startups, and online advertising. His goal with RealDirect is to use all the data he can access about real estate to improve the way people sell and buy houses.

Normally, people sell their homes about once every seven years, and they do so with the help of professional brokers and current data. But there’s a problem both with the broker system and the data quality. RealDirect addresses both of them.

First, the brokers. They are typically “free agents” operating on their own—think of them as home sales consultants. This means that they guard their data aggressively, and the really good ones have lots of experience. But in the grand scheme of things, that really means they have only slightly more data than the inexperienced brokers.

RealDirect is addressing this problem by hiring a team of licensed real estate agents who work together and pool their knowledge. To accomplish this, it built an interface for sellers, giving them useful data driven tips on how to sell their house. It also uses interaction data to give real-time recommendations on what to do next.

The team of brokers also become data experts, learning to use information-collecting tools to keep tabs on new and relevant data or to access publicly available information. For example, you can now get data on co-op (a certain kind of apartment in NYC) sales, but that’s a relatively recent change.

One problem with publicly available data is that it’s old news—there’s a three-month lag between a sale and when the data about that sale is available. RealDirect is working on real-time feeds on things like when people start searching for a home, what the initial offer is, the time between offer and close, and how people search for a home online. Ultimately, good information helps both the buyer and the seller. At least if they’re honest.

**How RealDirect Makes Money**

First, it offers a subscription to sellers—about $395 a month—to access the selling tools. Second, it allows sellers to use RealDirect’s agents at a reduced commission, typically 2% of the sale instead of the usual 2.5% or 3%. This is where the magic of data pooling comes in: it allows RealDirect to take a smaller commission because it’s more optimized, and therefore gets more volume.

The site itself is best thought of as a platform for buyers and sellers to manage their sale or purchase process. There are statuses for each person on site: active, offer made, offer rejected, showing, in contract, etc. Based on your status, different actions are suggested by the software.

Doug talked about key issues that a buyer might care about—nearby parks, subway, and schools, as well as the comparison of prices per square foot of apartments sold in the same building or block. This is the kind of data they want to increasingly cover as part of the service of RealDirect.

**Exercise**: Congrats your team just got hired as RealDirect’s dedicated Data Science consulting group.

Doug would like your team to work its’ data science magic to determine how the business model is functioning to help them invest strategically to ensure company growth.

Explore its existing website, (<https://www.realdirect.com/>) thinking about how buyers and sellers would navigate through it and try to understand the existing business model, and think about how analysis of RealDirect user-behavior data could be used to inform decision-making and product development. Come up with two research questions you think could be answered using data they have or could acquire that would help the company refine the business model. These can be oriented towards development of new features or analysis of the current platform.

For each research question also determine:

* The data you would advise the engineers log and what the dataset would look like, generally.
* The associated metric you would recommend the company track to know if the question you are proposing is being answer. (Target for Success)