Class 2 - Command Line

General Assembly – Data Science

Agenda for Tonight

- Heroku Exercise: Let's Make a Pie Chart
- Aurelion Geron: pgs. 3-8
- Test Installations
  - Anaconda
  - Atom
  - Python
  - Jupyter Notebooks
- Command Line
- Git & Github
  - Cloning a repo
  - Pushing a repo

What is Machine Learning

- Aurelion Geron: pgs. 3-8

[Machine Learning is the] field of study that gives computers the ability to learn without being explicitly programmed.

—Arthur Samuel, *1959*

A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.

—Tom Mitchell, *1997*

Email Spam Filter

your spam filter is a Machine Learning program that can learn to flag spam given examples of spam emails (e.g., flagged by users) and examples of regular (nonspam, also called “ham”) emails.

- The examples that the system uses to learn are called the *training set*.
- the task T is to flag spam for new emails
- the experience E is the *training data*
- the performance measure P needs to be defined (e.g., you can use the ratio of correctly classified emails)

Email Spam Filter

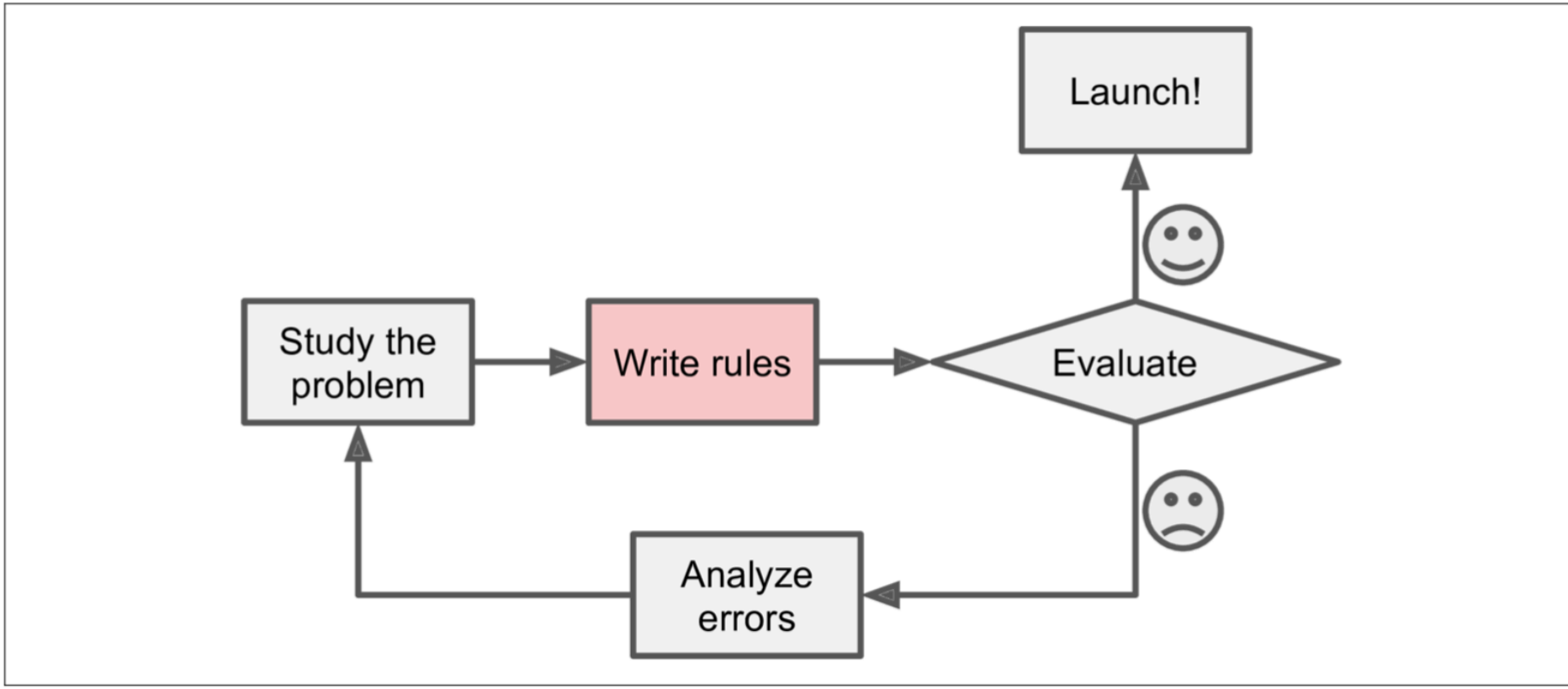

Figure 1-1. The traditional approach

Email Spam Filter

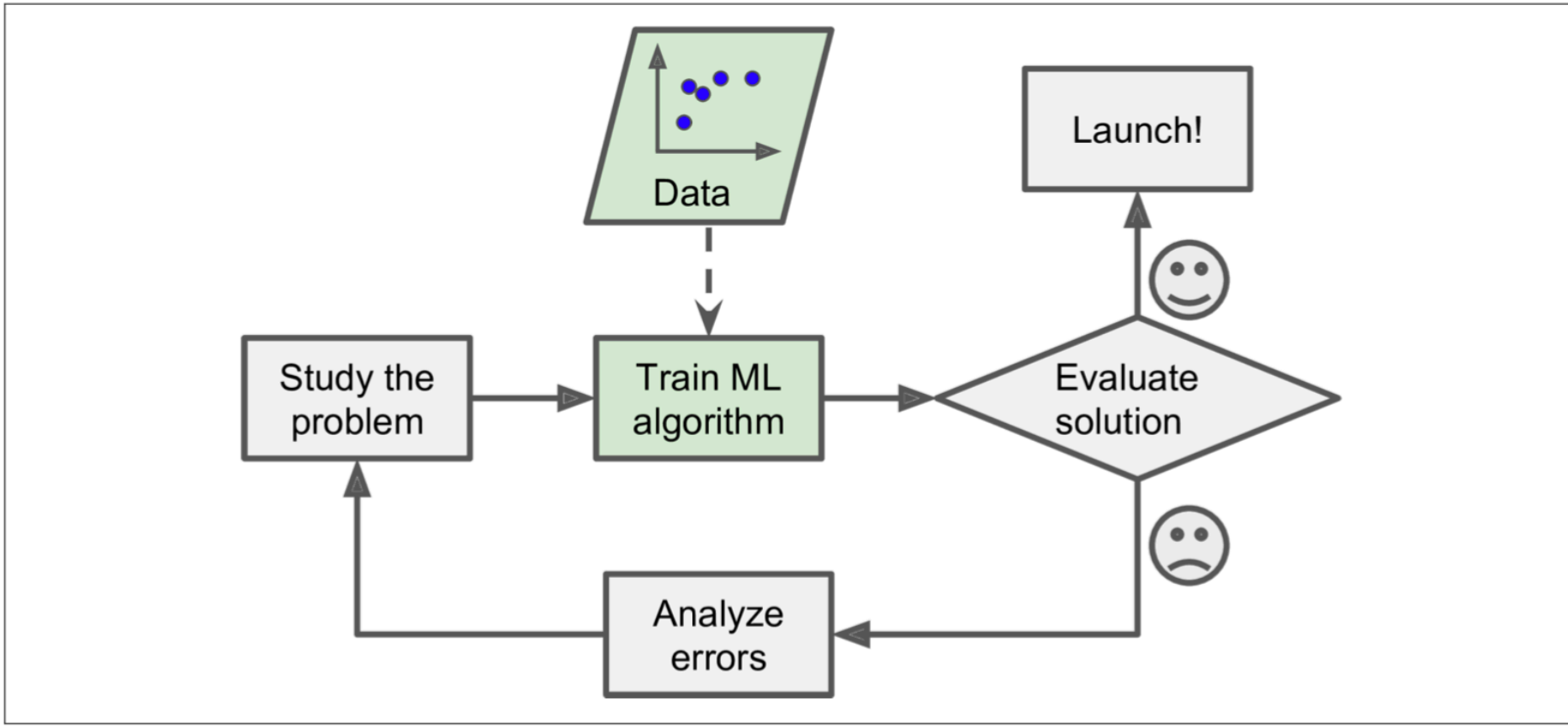

Figure 1-2. Machine Learning approach

Email Spam Filter

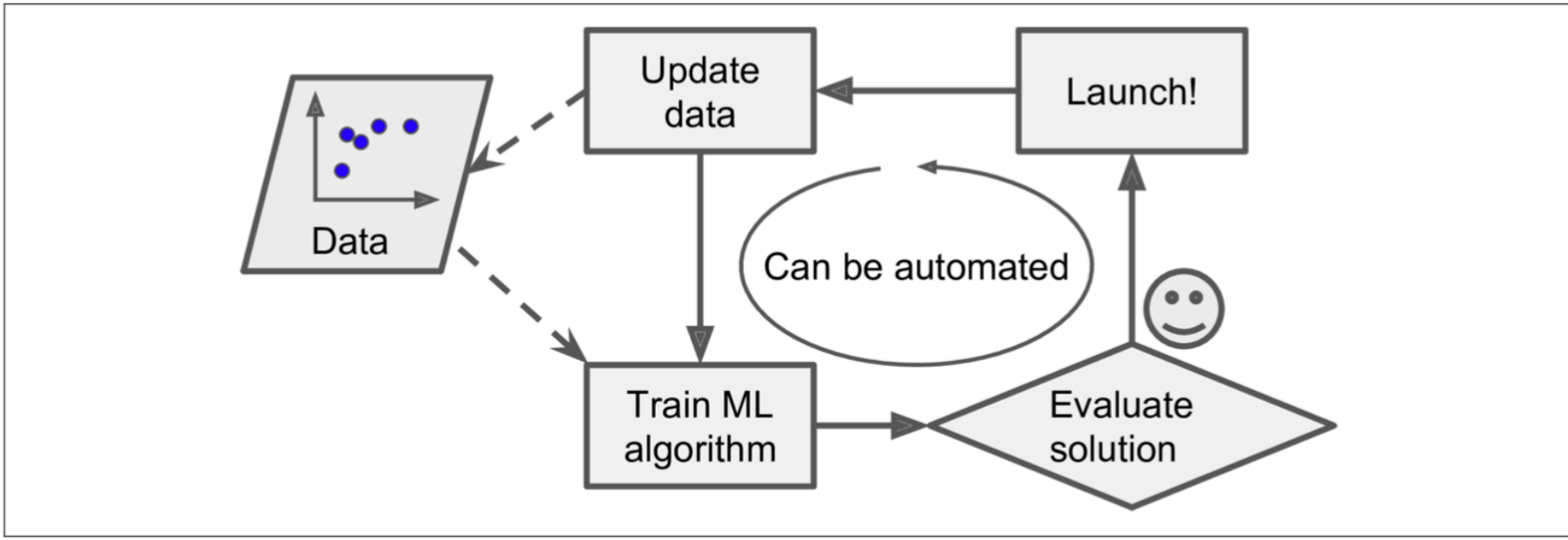

Figure 1-3. Automatically adapting to change

Email Spam Filter

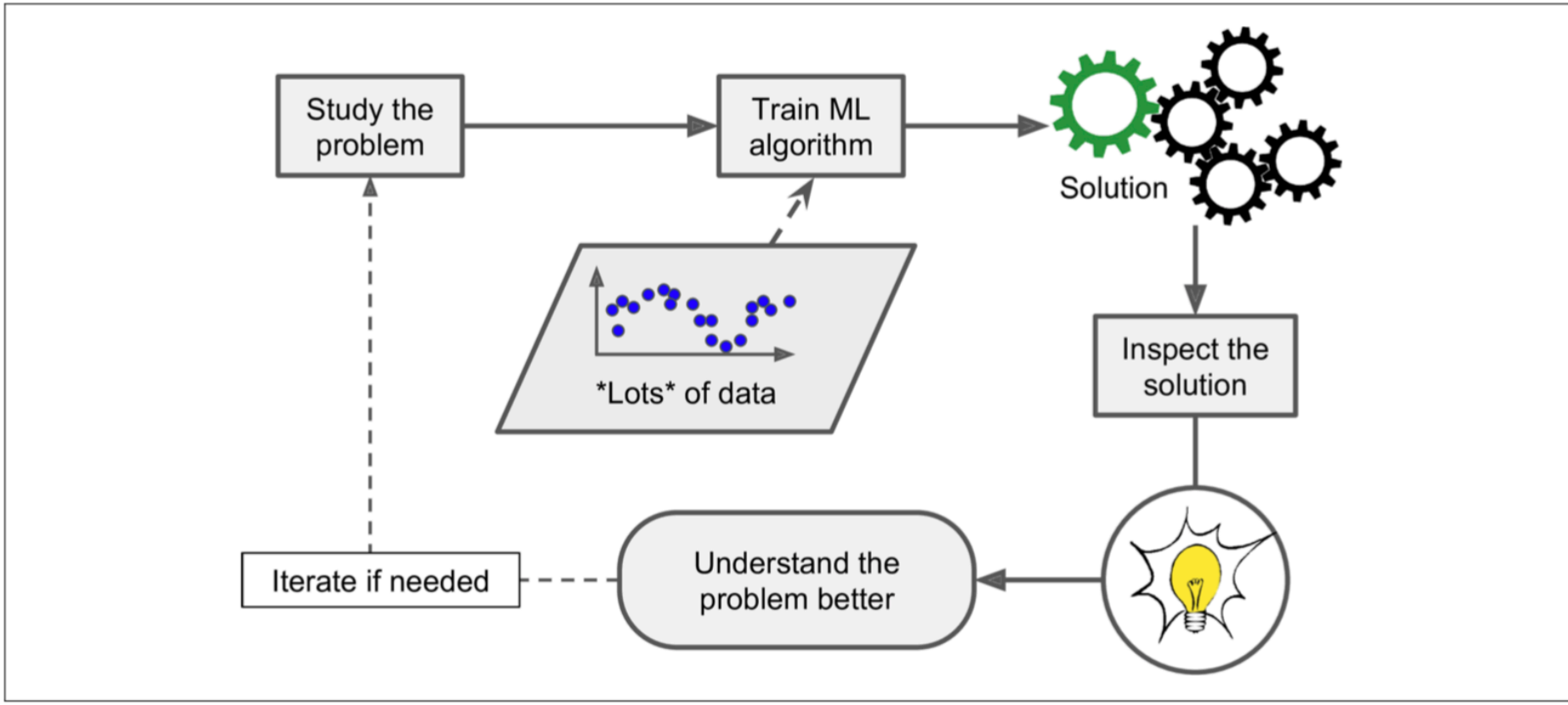

Figure 1-4. Machine Learning can help humans learn

Machine Learning is Great for:

- Problems for which existing solutions require a lot of hand-tuning or long lists of rules: one Machine Learning algorithm can often simplify code and perform better.
- Complex problems for which there is no good solution at all using a traditional approach: the best Machine Learning techniques can find a solution.
- Fluctuating environments: a Machine Learning system can adapt to new data.
- Getting insights about complex problems and large amounts of data.