Perspectives on Computational Research: Term Paper Re-write

Laurence Warner

lpwarner@uchicago.edu

Draft of Paper

Who on Earth doesn't own a Smartphone in 2018? Sociological predictions of technology uptake.

Abstract

Introduction

Smartphones are a technological device that has rapidly changed society in a unique way.

It is projected that by 2022, 90% of all mobile subscriptions will be for Internet-enabled “smartphones,” which are already in the majority (Ericsson, 2016). In the United States, more time is spent on digital activity on smartphones than on computers (ComScore, 2015).

Data

Source

Pew Research Center’s Internet Project Core Trends Survey, 2018. It can be accessed freely online here: http://www.pewinternet.org/datasets/

Interviews with a nationally representative sample of 2,002 adults were conducted between January 3-10 2018. The target population for the study is non-institutionalized persons age 18 and over, living in the US. Most of the interviews were conducted using cellphones (n = 1,502) with the remainder conducted using landlines (n=500); both groups were included in the final sample. According to the Pew Research Center, the landline sample was collected using a proportional sample based on listed telephone households. The cellphone sample was selected systematically from dedicated wireless numbers. This dataset contains questions about social media use in 2018 and attitudes toward the internet and whether Americans think it's good or bad for society. Random digit dialing was used to collect survey responses and the final sample was weighted to represent the American adult population. The sample response rate was 11%.

Cleaning

The dataset was cleaned using Python code, including the Pandas library. I chose to discard rows which contained answers including 'Don't Know' or 'Refused to answer'. This reduced the size of the dataset to n = 1567, which is easily enough to run machine learning models reliably (usually a minimum of around 500 is required).

In addition, in order to create a binary label:
1 = no smartphones
0 = smartphone
I grouped together folks without cell phones and folks with non-smart cell phones.

Literature Review

Content

The primary theoretical and empirical inspiration is Tsetsi et al. (2017): "Smartphone Internet access and use: Extending the digital divide and usage gap". They introduce the concept of the usage gap. This is an adaptation of the sociological theory of the knowledge gap, which posits that that knowledge, like other forms of wealth, is differentially distributed throughout a social system. So it states that technology usage is also distributed differentially.

They also use the Pew Internet survey - but from 2012. So I will be interested to compare results to see how demographic influences have changed in this fast-changing technological landscape.

Marler (2018): "Mobile phones and inequality: Findings, trends, and future directions" finds that socioeconomic effects are prevalent. But interestingly, low socioeconomic groups are often prone to having smartphones, but no other internet-devices - leading them to be "smartphone dependent". Is the income effect lower than we might expect then?

Andone et al. (2016) "How age and gender affect smartphone usage" use an alternative dataset: their own app "Menthal" to analyze smartphone usage. This data is not available - but I feel that survey data is superior because it represents the entire nation. Non-smartphone users are not captured by their smartphone app.

Methods & Results

Traditional Explanatory Analysis

Before exploring predictive machine learning models, to understand the dataset better, I firstly explore an explanatory analysis using traditional binary dependent variable models. I choose logistic regression as the most widely used and easily interpretable.

Table 1 below shows the regression output:

	coef	std err	z	P>|z|
sex	-2.7108	0.660	-4.105	0.000
age	0.0741	0.148	0.500	0.617
educ	0.9605	0.083	11.572	0.000
hisp	-0.2292	0.049	-4.707	0.000
inc	0.4006	0.272	1.470	0.141
race	-0.2983	0.035	-8.575	0.000

Conclusion and Future Work

References