|  |  |
| --- | --- |
| Algorithm: Ordinal Regression | |
| USN: 1MS18CS024 | NAME: Aravind S |
| USN: 1MS18CS029 | NAME: Arun J Kennedy |

**Description of the Algorithm: <<Write 2-3 Paragraphs about the Algorithm>>**

Ordinal Regression denotes a family of statistical learning methods in which the goal is to predict a variable which is discrete and ordered. For example, predicting the movie rating on a scale of 1 to 5 starts can be considered an ordinal regression task.

The logistic ordinal regression model, also known as the proportional odds was introduced in the early 80s by McCullagh and is a generalized linear model specially tailored for the case of predicting ordinal variables, that is, variables that are discrete (as in classification) but which can be ordered (as in regression). It can be seen as an extension of the logistic regression model to the ordinal setting.

**Algorithm Pseudocode:**

The following formulas must be implemented to achieve the Ordinal Regression.

Given X∈Rn×p input data and y∈Nn target values. For simplicity we assume y is a non-decreasing vector, that is, y1≤y2≤.... Just as the logistic regression models posterior probability P(y=j|Xi) as the logistic function, in the logistic ordinal regression we model the *cummulative* probability as the logistic function. That is,

P(y≤j|Xi)=ϕ(θj−wTXi)=11+exp(wTXi−θj)

where w,θ are vectors to be estimated from the data and ϕ is the logistic function defined as ϕ(t)=1/(1+exp(−t)).

Model estimation can be posed as an optimization problem. Here, we minimize the loss function for the model, defined as minus the log-likelihood:

L(w,θ)=−∑i=1nlog(ϕ(θyi−wTXi)−ϕ(θyi−1−wTXi))

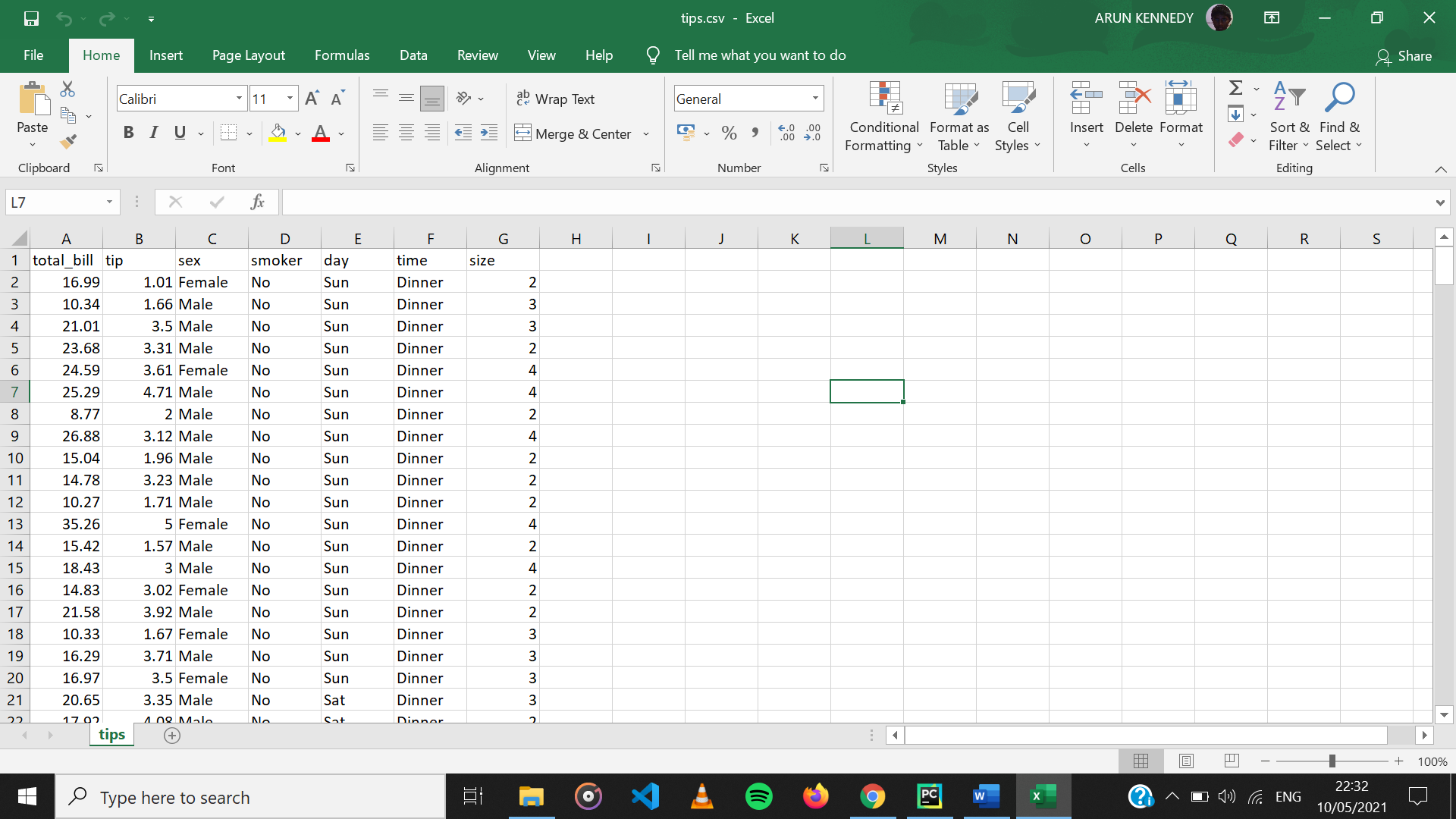
In this sum all terms are convex on w, thus the loss function is convex over w. It might be also jointly convex over w and θ L under the constraint that θ is a non-decreasing vector.

Using the formula log(ϕ(t))′=(1−ϕ(t)), we can compute the gradient of the loss function as

∇wL(w,θ)∇θL(w,θ)=∑i=1nXi(1−ϕ(θyi−wTXi)−ϕ(θyi−1−wTXi))=∑i=1neyi(1−ϕ(θyi−wTXi)−11−exp(θyi−1−θyi))+eyi−1(1−ϕ(θyi−1−wTXi)−11−exp(−(θyi−1−θyi)))(1)(2)(3)

where ei is the ith canonical vector.

**Data set Used: (Attach Screen shot of the few rows and also the Kaggle/Dataset link)**



Kaggle link: <https://www.kaggle.com/ranjeetjain3/seaborn-tips-dataset>

**Challenges faced during the implementation of the program:**

**Problem 1:** Numerical values are expected instead of categories.

**Solution:** Map the discrete categories to numerical values.

**Problem 2:** Not many references regarding Ordinal Regression were found.

**Solution 2:** Excessive searching for the resources.

**Advantages & Disadvantages of the Algorithm:**

|  |  |
| --- | --- |
| Advantages | Disadvantages |
| Logistic regression is easier to implement, interpret, and very efficient to train. | If the number of observations is lesser than the number of features, Logistic Regression should not be used, otherwise, it may lead to overfitting. |
| It makes no assumptions about distributions of classes in feature space. | It constructs linear boundaries. |
| It not only provides a measure of how appropriate a predictor (coefficient size) is, but also its direction of association (positive or negative). | It can only be used to predict discrete functions. Hence, the dependent variable of Logistic Regression is bound to the discrete number set. |
| It is very fast at classifying unknown records. | Non-linear problems can’t be solved with logistic regression because it has a linear decision surface. Linearly separable data is rarely found in real-world scenarios. |
| Good accuracy for many simple data sets and it performs well when the dataset is linearly separable. | Logistic Regression requires average or no multicollinearity between independent variables. |
| It not only provides a measure of how appropriate a predictor(coefficient size)is, but also its direction of association (positive or negative). | It can only be used to predict discrete functions. Hence, the dependent variable of Logistic Regression is bound to the discrete number set. |

**Applications of the Algorithm:**

* Investigating the factors that influence the sales of a particular product of a firm.
* Analyzing the factors and behaviors that influence the performance of an athlete.
* Investigate the likelihood of a person to join Grad-School.
* Determining the risk factors of child malnutrition.
* Identifying the determinants of an illness.

**Output: (Screen shots)**



**References:**

* <https://pythonhosted.org/mord/#:~:text=What%20is%20ordinal%20regression%20%3F,considered%20an%20ordinal%20regression%20task>.
* <http://fa.bianp.net/blog/2013/logistic-ordinal-regression/>
* <https://www.geeksforgeeks.org/advantages-and-disadvantages-of-logistic-regression/>
* <https://stats.idre.ucla.edu/r/dae/ordinal-logistic-regression/>