Predicting Food Health & Security

Victor Vulovic#1

#GalvanizeU, University of New Haven  
44 Tehama St, San Francisco, CA 94105, USA

1hello@victorvulovic.com

Abstract— This document gives an overview of data exploration, project procedure, and experimental results for 5 years worth of data from the Global Food Security Index (GFSI). Efforts on this dataset reflect a basic blueprint for other datasets of interest to “Project 8”, a cross-collaborative initiative of Nielsen, Inc. and The Demand Institute. Project 8 is currently working to create a massive global database dedicated to solving the 2020 Sustainable Development Goals of the U.N.

Keywords— GridSearch, Machine Learning, Global Development, Food Security, Public Policy

1. Introduction

In September of 2015, world leaders at the UN Summit declared 17 Sustainable Development Goals (SGDs) [1] to be achieved by 2030. These goals represent inspiration of all global citizens to collaborate on major world issues such as poverty, hunger, education, environmental protection, etc. Number two on the list of 17 is for Zero Hunger. This is what inspired a collaboration with Nielsen, The Demand Institute, The Conference Board, and the United Nations Foundation to create Project 8 [2].

As part of a Capstone assignment for GalvanizeU, I chose to work with Nielsen and The Demand Institute to attempt to model food security/risk of shortage. Known formally as Project 8, this current initiative aims to alleviate food insecurity through data awareness and analysis by aggregating the world’s key data sources that impact food security. The goal of this project was to create a model for assessing and predicting communities at risk of food insecurity.

Project 8 seeks to become a global data hub for human needs, a resource for policy-makers, development professionals, field-workers, and community leaders alike. As described by leaders at Project 8, there are 3 main attributes hindering progress in pursuit of organizing data on global human needs:

1. Fragmented: there is no shortage of data available, but rather there are very many different sources with many various collection, storage, and sharing methods. They are hosted by different organizations, countries, & communities and are usually not in the same format, language, and/or timeframe. Many countries lack any collection mechanism whatsoever.
2. Variable: despite universality in the nature of basic human needs, there is a litany of perspectives and methodologies for collecting and recording data on said needs. Comparing & contrasting different approaches to data collection is challenging, makes normalization difficult, and reduces consistency when attempting to query databases. Additionally, countless unwritten (and explicit) assumptions impede a universal understanding of data, and thus, how to most appropriately solve problems.
3. Cumbersome: actually accessing this data requires technical (data/software engineering) skills and knowledge, access to computing power, and enormous patience. No one source of data is exactly like the other. Many API’s rely upon differing technology stacks, have strict query limits, and sparse/confusing documentation.

The current approach to solving the issue is to host Project 8 as a social network and shared Salesforce CRM tool. The idea behind this is to have a well-designed and user-friendly interface for people to quickly and clearly get answers to any questions they may have regarding the 17 SGDs. So far, they have launched a beta prototype of the product and received feedback from early members.

Moving forward, the plans are to continue growing the user base, iterating the product & design, and creating value for current and future users. This last part is what my Capstone proposal intended to do for the initiative.

The following section outlines more of the specifics of the Capstone proposal and hypothesis for experiment & research.

1. Overview
2. *Hypothesis & Initial Proposal*
3. *Theory*

An easy way to comply with the conference paper formatting requirements is to use this document as a template and simply type your text into it.

1. Page Layout

Your paper must use a page size corresponding to A4 which is 210mm (8.27") wide and 297mm (11.69") long. The margins must be set as follows:

* Top = 19mm (0.75")
* Bottom = 43mm (1.69")
* Left = Right = 14.32mm (0.56")

Your paper must be in two column format with a space of 4.22mm (0.17") between columns.

1. Page Style

All paragraphs must be indented. All paragraphs must be justified, i.e. both left-justified and right-justified.

1. Text Font of Entire Document

The entire document should be in Times New Roman or Times font. Type 3 fonts must not be used. Other font types may be used if needed for special purposes.

Recommended font sizes are shown in Table 1.

1. Title and Author Details

Title must be in 24 pt Regular font. Author name must be in 11 pt Regular font. Author affiliation must be in 10 pt Italic. Email address must be in 9 pt Courier Regular font.

Model comparison  
GFSI - Health & Food Safety Prediction

|  |  |  |  |
| --- | --- | --- | --- |
| Mod. | Performance Metrics (1 year window) | | |
| Name | Accuracy | MSE |
| RF | Random Forest Classifer | 91.3% | 0.087 |
| KNN | K-Nearest Neighbors Classifier | 91.3% | 0.087 |
| SVC | Support Vector Classifier | 87.0% | 0.130 |
| LR | Logistic Regression | 82.6% | 0.174 |

All title and author details must be in single-column format and must be centered.

Every word in a title must be capitalized except for short minor words such as “a”, “an”, “and”, “as”, “at”, “by”, “for”, “from”, “if”, “in”, “into”, “on”, “or”, “of”, “the”, “to”, “with”.

Author details must not show any professional title (e.g. Managing Director), any academic title (e.g. Dr.) or any membership of any professional organization (e.g. Senior Member IEEE).

To avoid confusion, the family name must be written as the last part of each author name (e.g. John A.K. Smith).

Each affiliation must include, at the very least, the name of the company and the name of the country where the author is based (e.g. Causal Productions Pty Ltd, Australia).

Email address is compulsory for the corresponding author.

1. Section Headings

No more than 3 levels of headings should be used. All headings must be in 10pt font. Every word in a heading must be capitalized except for short minor words as listed in Section III-B.

1. Level-1 Heading: A level-1 heading must be in Small Caps, centered and numbered using uppercase Roman numerals. For example, see heading “III. Page Style” of this document. The two level-1 headings which must not be numbered are “Acknowledgment” and “References”.
2. Level-2 Heading: A level-2 heading must be in Italic, left-justified and numbered using an uppercase alphabetic letter followed by a period. For example, see heading “C. Section Headings” above.
3. *Level-3 Heading:* A level-3 heading must be indented, in Italic and numbered with an Arabic numeral followed by a right parenthesis. The level-3 heading must end with a colon. The body of the level-3 section immediately follows the level-3 heading in the same paragraph. For example, this paragraph begins with a level-3 heading.
4. Figures and Tables

Figures and tables must be centered in the column. Large figures and tables may span across both columns. Any table or figure that takes up more than 1 column width must be positioned either at the top or at the bottom of the page.

Graphics may be full color. All colors will be retained on the CDROM. Graphics must not use stipple fill patterns because they may not be reproduced properly. Please use only *SOLID FILL* colors which contrast well both on screen and on a black-and-white hardcopy, as shown in Fig. 1.

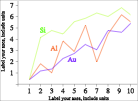


Fig. 1 A sample line graph using colors which contrast well both on screen and on a black-and-white hardcopy

Fig. 2 shows an example of a low-resolution image which would not be acceptable, whereas Fig. 3 shows an example of an image with adequate resolution. Check that the resolution is adequate to reveal the important detail in the figure.

Please check all figures in your paper both on screen and on a black-and-white hardcopy. When you check your paper on a black-and-white hardcopy, please ensure that:

* the colors used in each figure contrast well,
* the image used in each figure is clear,
* all text labels in each figure are legible.

1. Figure Captions

Figures must be numbered using Arabic numerals. Figure captions must be in 8 pt Regular font. Captions of a single line (e.g. Fig. 2) must be centered whereas multi-line captions must be justified (e.g. Fig. 1). Captions with figure numbers must be placed after their associated figures, as shown in Fig. 1.



Fig. 2 Example of an unacceptable low-resolution image



Fig. 3 Example of an image with acceptable resolution

1. Table Captions

Tables must be numbered using uppercase Roman numerals. Table captions must be centred and in 8 pt Regular font with Small Caps. Every word in a table caption must be capitalized except for short minor words as listed in Section III-B. Captions with table numbers must be placed before their associated tables, as shown in Table 1.

1. Page Numbers, Headers and Footers

Page numbers, headers and footers must not be used.

1. Links and Bookmarks

All hypertext links and section bookmarks will be removed from papers during the processing of papers for publication. If you need to refer to an Internet email address or URL in your paper, you must type out the address or URL fully in Regular font.

1. References

The heading of the References section must not be numbered. All reference items must be in 8 pt font. Please use Regular and Italic styles to distinguish different fields as shown in the References section. Number the reference items consecutively in square brackets (e.g. [1]).

When referring to a reference item, please simply use the reference number, as in [2]. Do not use “Ref. [3]” or “Reference [3]” except at the beginning of a sentence, e.g. “Reference [3] shows …”. Multiple references are each numbered with separate brackets (e.g. [2], [3], [4]–[6]).

Examples of reference items of different categories shown in the References section include:

* example of a book in [1]
* example of a book in a series in [2]
* example of a journal article in [3]
* example of a conference paper in [4]
* example of a patent in [5]
* example of a website in [6]
* example of a web page in [7]
* example of a databook as a manual in [8]
* example of a datasheet in [9]
* example of a master’s thesis in [10]
* example of a technical report in [11]
* example of a standard in [12]

1. Conclusions

The version of this template is V2. Most of the formatting instructions in this document have been compiled by Causal Productions from the IEEE LaTeX style files. Causal Productions offers both A4 templates and US Letter templates for LaTeX and Microsoft Word. The LaTeX templates depend on the official IEEEtran.cls and IEEEtran.bst files, whereas the Microsoft Word templates are self-contained. Causal Productions has used its best efforts to ensure that the templates have the same appearance.

Causal Productions permits the distribution and revision of these templates on the condition that Causal Productions is credited in the revised template as follows: “original version of this template was provided by courtesy of Causal Productions (www.causalproductions.com)”.

Acknowledgment

The heading of the Acknowledgment section and the References section must not be numbered.

Causal Productions wishes to acknowledge Michael Shell and other contributors for developing and maintaining the IEEE LaTeX style files which have been used in the preparation of this template. To see the list of contributors, please refer to the top of file IEEETran.cls in the IEEE LaTeX distribution.

References

1. (2017) The United Nations Website. [Online]. Available: <http://www.un.org/sustainabledevelopment/sustainable-development-goals/>
2. (2017) The Project 8 Website. [Online]. Available: http://demandinstitute.org/projects/project-8/
3. J. Breckling, Ed., *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.
4. S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok, “A novel ultrathin elevated channel low-temperature poly-Si TFT,” *IEEE Electron Device Lett.*, vol. 20, pp. 569–571, Nov. 1999.
5. M. Wegmuller, J. P. von der Weid, P. Oberson, and N. Gisin, “High resolution fiber distributed measurements with coherent OFDR,” in *Proc. ECOC’00*, 2000, paper 11.3.4, p. 109.
6. R. E. Sorace, V. S. Reinhardt, and S. A. Vaughn, “High-speed digital-to-RF converter,” U.S. Patent 5 668 842, Sept. 16, 1997.
7. (2002) The IEEE website. [Online]. Available: http://www.ieee.org/
8. M. Shell. (2002) IEEEtran homepage on CTAN. [Online]. Available: http://www.ctan.org/tex-archive/macros/latex/contrib/supported/IEEEtran/
9. *FLEXChip Signal Processor (MC68175/D)*, Motorola, 1996.
10. “PDCA12-70 data sheet,” Opto Speed SA, Mezzovico, Switzerland.
11. A. Karnik, “Performance of TCP congestion control with rate feedback: TCP/ABR and rate adaptive TCP/IP,” M. Eng. thesis, Indian Institute of Science, Bangalore, India, Jan. 1999.
12. J. Padhye, V. Firoiu, and D. Towsley, “A stochastic model of TCP Reno congestion avoidance and control,” Univ. of Massachusetts, Amherst, MA, CMPSCI Tech. Rep. 99-02, 1999.
13. *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification*, IEEE Std. 802.11, 1997.