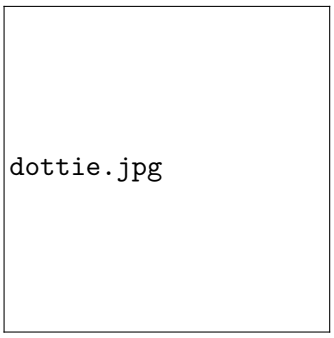
SOMM: INTO THE MODEL

— AI AND DATA SCIENCE FOR WINE PROFESSIONALS

SHENGLI HU. PHD.

SHENGLIHU1224@GMAIL.COM

About the Author



Shengli Hu received her PhD from Cornell University in 2019, and has since been an AI research scientist with Dataminr Inc. in New York City. Her research experience and interests lie in language and vision, in the context of marketing, behavioral economics, and information management, with publications in top conferences and journals in natural language processing, computer vision, and applied statistics including Association of Computational Linguistics (ACL), Empirical Methods in Natural Language Processing (EMNLP), European Conference on Computer Vision (ECCV), Computer Vision and Pattern Recognition (CVPR) and Annals of Applied Statistics (AoAS).

She is also a wine professional with credentials including Certified Sommelier by The Court of Master Sommelier, Certified Specialist of Wine by The Society of Wine Educators, Certified Specialist of Spirits by The Society of Wine Educators, and Level 4 Diploma by Wine Spirits Education Trust. She is an incoming Master of Wine student.

To wine professionals and enthusiasts curious about AI and data science, as well as data scientists and AI research scientists who caught the wine bug as the author did.

Table of Contents

1 What is a Sommelier? 5

1.1 Theory . . . 5

1.1.1 Knowledge Graph . . . 5

1.1.2 Questions Answering . . . 10

1.1.3 Cross-lingual Embedding . . . 17

1.2 Deductive Tasting . . . 23

1.2.1 Decision Tree . . . 23

1.2.2 Structured Prediction . . . 31

1.3 Service . . . 39

1.3.1 Wine and Food Pairing: Matching Networks . . . 39

1.3.2 Recommendation Systems . . . 44

1.3.3 Cocktails: Semantic Networks . . . 50

1.4 Evaluation . . . 56

1.4.1 Detecting Credibility and Expertise in Text . . . 57

1.4.2 Detecting Credibility and Expertise in Speech . . . 61

1.4.3 Multimodal Learning . . . 65

1.4.4 Multitask Learning . . . 71

2 What is Natural Wine? 78

2.1 Terrior: Causal Inference with Natural Experiments . . . 80

2.2 Terrior: Causal Inference with Regression Discontinuity . . . 85

2.3 Does Natural Wine Taste Different? . . . 92

2.3.1 Topic Modeling . . . 93

2.3.2 Permutation Tests . . . 100

2.3.3 Propensity Score Matching . . . 104

3 Wine Laws and Regulations 109

3.1 Wine Labels . . . 110

3.1.1 Optical Character Recognition . . . 111

3.1.2 Object Detection . . . 113

3.1.3 Image Retrieval . . . 123

3.1.4 Style Transfer . . . 130

3.2 Wine Tariffs . . . 138

3.2.1 Causal Inference with Difference-in-Difference . . . 143

3.2.2 Social Voting . . . 149

4 Have You Heard About The Cheating Scandal? 156

4.1 Detecting Information Concealment . . . 156

4.2 Deception Detection . . . 166

4.2.1 Deception Detection in Speech . . . 167

4.2.2 Deception Detection in Text . . . 172

4.2.3 Deception Detection in Other Modalities 177

5 What Makes a Great Wine List? 180

5.1 Semantic Parsing for Wine Lists 180

5.2 Automatic Wine List Evaluation 185

5.3 Automatic Wine List Generation 191

19 References 198