The State of the Art in Enhancing Trust in Machine Learning Models with the Use of Visualizations: Survey Summary

Table 1: The summary of our survey (as of March 17, 2020). See the updated survey data at `https://trustmlvis.lnu.se`

| Technique | A.1.1 Biology | A.1.2 Business | A.1.3 Computer Vision | A.1.4 Computer | A.1.5 Health | A.1.6 Humanities | A.1.7 Industry | A.1.8 Simulation | A.1.9 Social / Socioeconomic | A.1.10 Other (Domain) | A.2.1 Binary | A.2.2 Multi-class | A.2.3 Multi-field | A.2.4 Other (Target Var.) | B.1.1 Convolutional Neural Network (CNN) | B.1.2 Deep Convolutional Network (DCN) | B.1.3 Deep Feed Forward (DFF) | B.1.4 Deep Neural Network (DNN) | B.1.5 Deep Q-Network (DQN) | B.1.6 Generative Adversarial Network (GAN) | B.1.7 Long Short-Term Memory (LSTM) | B.1.8 Recurrent Neural Network (RNN) | B.1.9 Variational Autoencoder (VAE) | B.1.10 Other (DL Methods) | B.2.1 Linear | B.2.2 Non-linear | B.3.1 Bagging | B.3.2 Boosting | B.3.3 Stacking | B.4.1 Other (Generic ML Methods) | B.2.1 Classification | B.2.2 Regression | B.2.3 Other | B.2.1 Association | B.2.2 Clustering | B.2.3 Dimr. Reduction | B.2.4 DR | B.2.1 Classification | B.2.2 Clustering | B.2.3 Classification | B.2.4 Control | C.1 Pre-processing / Input | C.2 In-processing / Model | C.3 Post-processing / Output | D.1 Model-agnostic / Black Box | D.2 Model-specific / White Box | E.1.1 2D | E.1.2 3D | E.2.1 Computed | E.2.2 Mapped | E.2.3 Abstract/Semi-Abstract Information | E.2.4 Individual / Individual | E.3.1 Bar Charts | E.3.2 Box Plots | E.3.3 Matrix | E.4.1 Glyphs / Icons / Thumbnails | E.4.2 Glyph-based Approaches | E.4.3 Heatmaps | E.4.4 Pixel-based Approaches | E.4.5 Tables | E.4.6 Text Clouds | E.4.7 Other | E.4.8 Line Plots | E.4.9 Line Charts | E.4.10 Node-link Diagrams | E.4.11 Parallel Coordinates Plots (PCP) | E.4.12 Flow-based Approaches | E.4.13 Radial Layouts | E.4.14 Scatterplots / Scatterplots Matrices (SPLOMs) | E.4.15 Scatterplots / Projection | E.4.16 Similarity Layouts | E.4.17 Tables / Lists | E.4.18 Treemap | E.4.19 Other (Visual.) | E.5.1 Select | E.5.2 Explore / Browse | E.5.3 Reconfigure | E.5.4 Encode | E.5.5 Filter / Query | E.5.6 Abstract / Elaborate | E.5.7 Connect | E.5.8 Guide / Shepard | E.5.9 Verbalize | E.6.1 Color | E.6.2 Opacity | E.6.3 Position / Orientation | E.6.4 Shape | E.6.5 Size | E.6.6 Texture | F.1.1 Standard | F.1.2 Comparison | F.2.1 Iterative Design / Development | F.2.2 After Development | F.3 Not Evaluated | G.1.1 Source Reliability | G.1.2 Transparent Collection Process | G.1.3 Uncertainty Awareness | G.2.1 Bias | G.2.2 Equality / Bias Bias | G.2.3 Fairness / Non-Discrimination | G.2.4 Guidance / Recommendation | G.2.5 Outlier Detection | G.3.1 Familiarity | G.3.2.1 Understanding / Explanation | G.3.2.2 Debugging / Diagnosis | G.3.2.3 Refinement / Steering | G.3.2.4 Comparison | G.3.3 Trustworthiness | G.3.4 Fairness | G.4.1 Experience | G.4.2 In Situ Comparison | G.4.3 Performance | G.4.4 What-if Hypotheses | G.4.5 Model Bias | G.4.6 Model Trustworthiness | G.5.1 Application of Colleagues | G.5.2 Visualization Evaluation | G.5.3 Metrics Validation / Benefits | G.5.4 User Bias | H.1 Practitioners / Domain Experts | H.2 Beginners | H.3 Designers | H.4 ML Experts |
|---|---|---

Table 2: The summary of our survey (continued).

| Technique | A.1.1 Biology | A.1.2 Business | A.1.3 Computer Vision | A.1.4 Computers | A.1.5 Health | A.1.6 Humanities | A.1.7 Manufacturing | A.1.8 Social | A.1.9 Social / Socioeconomic | A.1.10 Other (Domain) | A.2.1 Binary | A.2.2 Multi-class | A.2.3 Multi-label | A.2.4 Continuous | A.2.5 Other (Output Var.) | B.1.1 Convolutional Neural Network (CNN) | B.1.2 Deep Convolutional Network (DCN) | B.1.3 Deep Feed Forward (DFF) | B.1.4 Deep Neural Network (DNN) | B.1.5 Deep Q-Network (DQN) | B.1.6 Generative Adversarial Network (GAN) | B.1.7 Graph Neural Network (GNN) | B.1.8 Recurrent Neural Network (RNN) | B.1.9 Variational Auto-Encoder (VAE) | B.1.10 Other (DL Methods) | B.2.1 Linear | B.2.2 Nonlinear | B.3.1 Bagging | B.3.2 Boosting | B.3.3 Stacking | B.3.4 Other (Generic ML Methods) | B.2.1.1 Classification | B.2.1.2 Regression | B.2.1.3 Other | B.2.2 Association | B.2.2.1 Clustering | B.2.3 Other | B.2.3.1 Classification | B.2.3.2 Clustering | B.2.4.1 Classification | B.2.4.2 Control | C.1 Preprocessing / Input | C.2 Preprocessing / Hyper | C.3 Model Ensemble | C.4 Model Selection | D.1 Model-agnostic / Black Box | D.2 Model-specific / White Box | E.1.1 2D | E.1.2 3D | E.2.1 Computed | E.2.2 Mapped | E.3.1 Saliency / Attribution | E.3.2 Neighbored Information | E.3.3 Instance / Instance / Influential | E.4.1 Bar Charts | E.4.2 Box Plots | E.4.3 Matrix | E.4.4 Glyphs / Icons / Thumbnails | E.4.5 Grid-based Approaches | E.4.6 Hierarchy | E.4.7 Hierarchy | E.4.8 Hierarchy | E.4.9 Line Charts | E.4.10 Node-link Diagrams | E.4.11 Parallel Coordinates Plots (PCP) | E.4.12 Pixel-based Approaches | E.4.13 Radial Layouts | E.4.14 Scatterplots / Projections | E.4.15 Scatterplots / Projections | E.4.16 Similarity Layouts | E.4.17 Tables / Lists | E.4.18 Treemaps | E.4.19 Other (Visual) | E.5.1 Select | E.5.2 Explore / Browse | E.5.3 Reconfigure | E.5.4 Encode | E.5.5 Filter / Query | E.5.6 Abstract / Elaborate | E.5.7 Connect | E.5.8 Guide / Shepherd | E.5.9 Verbalize | E.6.1 Color | E.6.2 Gestalt | E.6.3 Position / Orientation | E.6.4 Shape | E.6.5 Size | E.6.6 Texture | F.1.1 Standard | F.1.2 Comparison | F.1.3 Design Development | F.2.1 After Development | F.2.2 After Development | F.3 Not Evaluated | G.1.1 Source Reliability | G.1.2 Transparent Collection Process | G.2.1 Uncertainty Awareness | G.2.2 Equality / Data Bias | G.2.3 Consumer Bias (Preferences) | G.2.4 Consumer Bias (Communication | G.2.5 Outlier Detection | G.3.1 Familiarity | G.3.2.1 Understanding / Explanation | G.3.2.2 Debugging / Diagnosis | G.3.2.3 Refinement / Steering | G.3.3.1 Comparison | G.3.3.2 Verifiability | G.3.4 Patterns | G.4.1 Experience | G.4.2 In Situ Comparison | G.4.3 Performance | G.4.4 What-if Hypotheses | G.4.5 Model Bias | G.4.6 Model Trust | G.5.1 Accountant of Colleagues | G.5.2 Visualization Evaluation | G.5.3 Metrics Validation / Results | G.5.4 User Bias | H.1 Beginners | H.2 Practitioners / Domain Experts | H.3 Developers | H.4 ML Experts |
|---|---|---