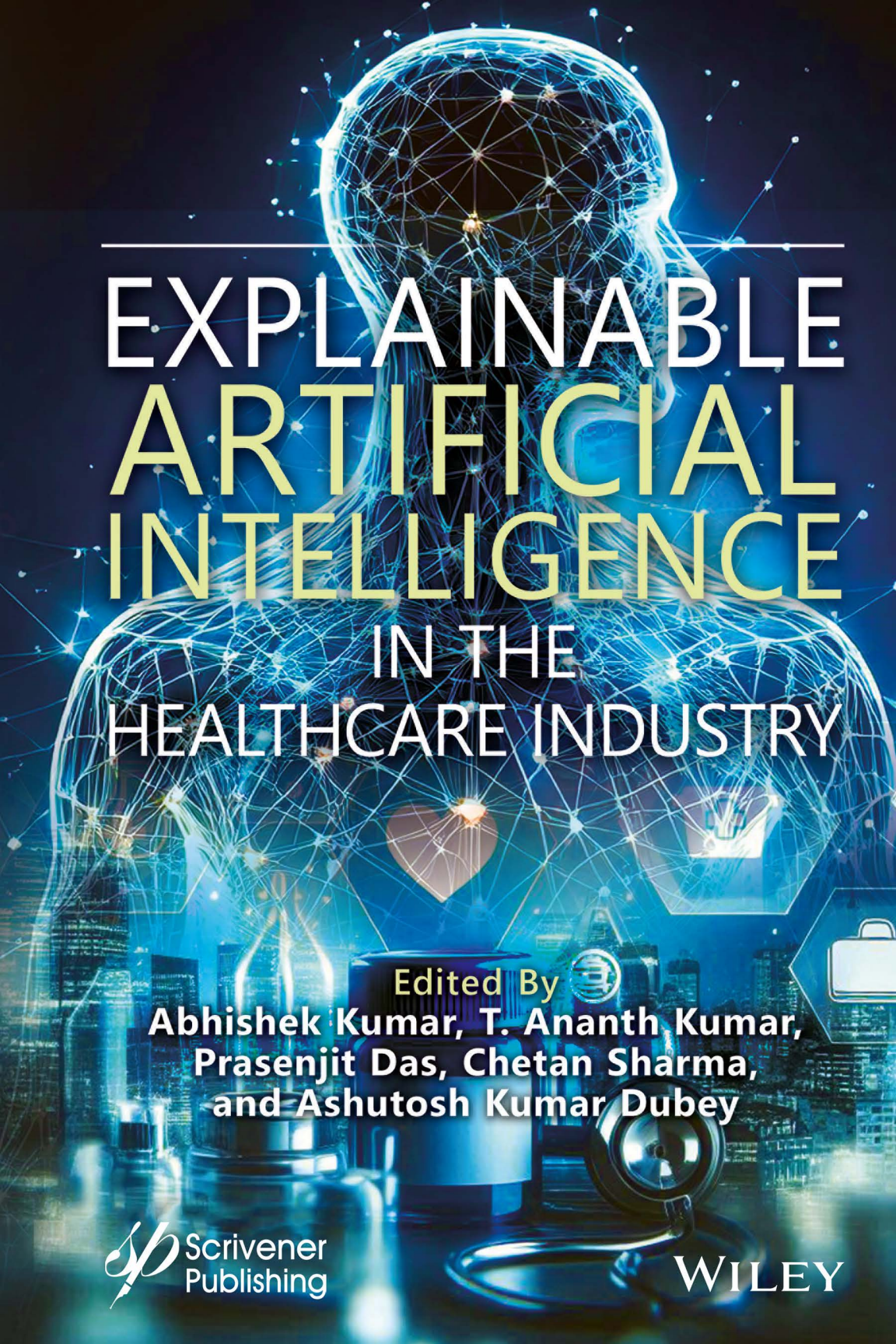




EXPLAINABLE ARTIFICIAL INTELLIGENCE IN THE HEALTHCARE INDUSTRY

Edited By

**Abhishek Kumar, T. Ananth Kumar,
Prasenjit Das, Chetan Sharma,
and Ashutosh Kumar Dubey**

 **Scrivener
Publishing**

WILEY

Explainable Artificial Intelligence in the Healthcare Industry

Scrivener Publishing

100 Cummings Center, Suite 541J
Beverly, MA 01915-6106

Publishers at Scrivener

Martin Scrivener (martin@scrivenerpublishing.com)
Phillip Carmical (pcarmical@scrivenerpublishing.com)

Explainable Artificial Intelligence in the Healthcare Industry

Edited by

Abhishek Kumar

T. Ananth Kumar

Prasenjit Das

Chetan Sharma

and

Ashutosh Kumar Dubey



WILEY

This edition first published 2025 by John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA and Scrivener Publishing LLC, 100 Cummings Center, Suite 541J, Beverly, MA 01915, USA

© 2025 Scrivener Publishing LLC

For more information about Scrivener publications please visit www.scrivenerpublishing.com.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

Wiley Global Headquarters

111 River Street, Hoboken, NJ 07030, USA

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Limit of Liability/Disclaimer of Warranty

While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials, or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read.

Library of Congress Cataloging-in-Publication Data

ISBN 9781394249268

Front cover images supplied by Adobe Firefly

Cover design by Russell Richardson

Set in size of 11pt and Minion Pro by Manila Typesetting Company, Makati, Philippines

Printed in the USA

10 9 8 7 6 5 4 3 2 1

Contents

Preface	xxix
1 A Review on Explainable Artificial Intelligence for Healthcare	1
<i>Rakhi Chauhan</i>	
1.1 Introduction	2
1.2 Literature Review	3
1.3 Reason for Using XAI	7
1.3.1 Applications of XAI Algorithms in Different Scenarios	8
1.3.2 Discussion on Different Aspects of XAI	8
1.4 Challenges and Future Prospects	10
1.5 Conclusion	11
References	12
2 Explainable Artificial Intelligence (XAI) in Healthcare: Fostering Transparency, Accountability, and Responsible AI Deployment	17
<i>Asha S. Manek, Shruti Vashist, Geeta Tripathi and Savita Sindhu</i>	
2.1 Introduction	18
2.2 Roles of XAI	20
2.3 Why XAI	22
2.4 XAI in Different Sectors	23
2.5 XAI in Healthcare	26
2.6 Challenges of XAI Adoption	28
2.7 Models Used in XAI Adoption	30
2.8 Conclusion	32
2.9 Future of XAI in Healthcare	33
References	34

3	Illuminating the Diagnostic Path: Unveiling Explainability in Medical Imaging	39
	<i>Sivanantham S., Anwar Basha H., Thanuja K., Shafiya Banu M., Maithili K. and AnilKumar Ambore</i>	
3.1	Introduction	40
3.2	The Need for Explainability in Medical Imaging AI	42
3.3	Related Works	43
3.4	Explainable AI Techniques for Medical Imaging	44
3.4.1	Saliency Maps	44
3.4.2	Attention Mechanisms	46
3.4.3	Decision Trees	46
3.4.4	Gradient-Weighted Class Activation Mapping (Grad-CAM)	47
3.4.5	LIME (Local Interpretable Model—Agnostic Explanations)	47
3.4.6	Integrated Gradients	48
3.4.7	SHAP (SHapley Additive exPlanations)	48
3.4.8	Layer-Wise Relevance Propagation (LRP)	48
3.4.9	SmoothGrad	49
3.4.10	Occlusion Sensitivity	49
3.4.11	Feature Visualization	50
3.4.12	Attention Condensation	51
3.4.13	Counterfactual Explanations	51
3.5	Real-World Applications and Case Studies	51
3.6	Deep Learning Approaches for Pneumonia Identification in Chest X-Ray Images: Methods and Methodology	53
3.6.1	Dataset Collection and Preprocessing	53
3.6.2	Deep Learning Model Selection	53
3.6.3	VGG16	53
3.6.4	Inception V3	54
3.6.5	ResNet 50V2	55
3.7	Model Training and Evaluation	55
3.8	Performance Evaluation	58
3.9	Comparative Analysis	59
3.9.1	Implications for Regulatory Compliance, Patient Safety, and Ethics	60
3.9.2	Regulatory Compliance	60
3.9.3	Patient Safety	60
3.10	Ethical Considerations	61

3.11	Advancing Clinical Workflows with Explainable AI	61
3.12	Conclusion	62
	References	63
4	HealsHealthAI: Unveiling Personalized Healthcare Insights with Open Source Fine-Tuned LLM	67
	<i>Lavan J. V. and Lakshmi Sangeetha</i>	
4.1	Introduction	68
4.1.1	HealsHealthAI Objective	68
4.2	Review Analysis	68
4.2.1	Literature Review	68
4.2.2	Additional Review	69
4.3	Motivation	70
4.3.1	Motto	70
4.4	Methodology	70
4.4.1	Data Collection and Processing	70
4.4.2	Storage	70
4.4.3	User Interaction	71
4.4.4	Model Creation and Loading	71
4.5	How LLMs, FAISS, and Langchain are Utilized in HealsHealthAI	71
4.5.1	Personalized Insights	71
4.5.2	Langchain in HealsHealthAI	71
4.6	Model Overview and Working	72
4.6.1	Overall Flowchart	72
4.6.2	Showcases of HealsHealthAI's Functionality	73
4.7	Community Contribution	73
4.7.1	The Call for Contributions	73
4.8	Experimental Findings	74
4.8.1	Quantitative Performance Metrics	74
4.8.2	Results	75
4.8.3	Scalability and Resource Usage	75
4.9	Challenges and Future Paths	76
4.9.1	Known Issues	76
4.9.2	Deployment Without Internet Connectivity	76
4.10	Conclusion	76
	References	76

5	Introduction to Explainable AI in EEG Signal Processing: A Review	79
	<i>Parag Puranik and Rahul Pethe</i>	
5.1	Introduction	80
5.2	Solution Approaches	82
5.2.1	Survey of Machine Learning for EEG Signal Processing	82
5.2.2	DL Procedures for EEG Signal Processing	85
5.2.2.1	Neural Network for EEG	87
5.2.3	Feature Extraction and Pre-Processing for EE	87
5.3	Classification Models for EEG Signals	87
5.4	Results and Discussions	93
5.5	Future Scope and Issues	95
5.6	Proposed Methodology	96
5.7	Conclusions	98
	References	99
6	Transparency in Disease Diagnosis: Leveraging Interpretable Machine Learning in Healthcare	105
	<i>Inam Ul Haq, Adil Husain Rather, Syed Zoofa Rufai, Ahmad Shah, Sheetal and Akib Mohi Ud Din Khanday</i>	
6.1	Introduction	106
6.2	Introduction to the Overarching Theme of Transparency and Accountability in AI-Driven Healthcare Solutions	108
6.3	The Significance of Model Interpretability in Clinical Decision Making and Patient Management Within the Context of Disease Diagnosis	110
6.4	Importance of Model Interpretability	112
6.5	Examination of How Interpretable ML Models Facilitate Clearer Understanding and Interpretation of Diagnostic Decisions	113
6.6	Importance of Transparency in Ensuring Accountability and Fostering Acceptance of AI Technologies Among Healthcare Professionals and Patients	115
6.7	Techniques for Interpretable Machine Learning	116
6.8	Explanation of How Each Technique Contributes to Enhancing Diagnostic Accuracy While Providing Transparent Decision Rationale	118
6.9	Comparative Analysis of Different Interpretable ML Approaches in Terms of Their Applicability and Effectiveness in Disease Diagnosis	119
6.10	Case Studies and Applications	121

6.11	Examination of How Interpretable ML has been Utilized to Improve Disease Diagnosis, Treatment Planning, and Patient Outcomes	124
6.12	Analysis of the Benefits and Challenges Associated with Implementing Interpretable ML in Clinical Practice Through Case Examples and Empirical Evidence	125
6.13	Conclusion	126
	References	127
7	Transparency in Text: Unraveling Explainability in Healthcare Natural Language Processing	131
	<i>Madhan Veeramani, Karthick P., S. Venkateswaran, Sriman B., Shaik Thasleem Bhanu and V. Seedha Devi</i>	
7.1	Introduction	132
7.2	Research Objectives	134
7.2.1	The Promise of NLP in Healthcare	135
7.2.2	Challenges in NLP Models in Healthcare	135
7.2.3	The Importance of Transparency and Explainability	136
7.3	Current Research Landscape	136
7.4	Related Works	137
7.5	Literature Survey	139
7.6	The Role of Explainability in Healthcare NLP	141
7.6.1	Defining Explainability in Healthcare NLP	141
7.6.2	Limitations of Traditional Black-Box NLP Models	141
7.6.3	Benefits of Transparent and Interpretable NLP Models	142
7.7	Enhancing Clinical Documentation with Interpretable NLP	143
7.7.1	Methodology	145
7.7.2	System Development	146
7.7.3	Integration of Explainable AI Techniques	147
7.7.3.1	Steps	148
7.8	Literature Survey: Integrating Explainable AI Techniques Into Natural Language Processing for Healthcare	149
7.9	Implementation Process	151
7.10	Results	153
7.11	Challenges and Future Directions	154
7.12	Future Directions	154
7.13	Ethical Considerations and Regulatory Implications	155
7.14	Conclusion	155
	References	156

8	Introduction to Explainable AI in Healthcare: Enhancing Transparency and Trust	161
	<i>Karthik Srinivasan, Chaithanya Kumar Viralam Ramamurthy, Saravanan Matheswaran and Shermin Shamsudheen</i>	
8.1	Introduction to Explainable AI in Healthcare	162
8.1.1	The Evolution of AI in Healthcare	162
8.1.2	The Need for Explainable AI	162
8.2	Exploring Explainability Techniques	163
8.2.1	Machine Learning Explainability Methods	164
8.2.2	Explainability in Deep Learning	164
8.2.3	Explainable Natural Language Processing (NLP)	165
8.3	Real-World Applications and Case Studies	166
8.3.1	Enhancing Clinician Comprehension	166
8.3.2	Facilitating Regulatory Compliance	167
8.3.3	Fostering Patient Trust	168
8.3.4	Patient-Centric Applications of Explainable AI	168
8.3.4.1	Shared Decision Making	168
8.3.4.2	Personalized Care Planning	169
8.3.4.3	Empowering Patient Advocacy	169
8.4	Challenges and Ethical Considerations	169
8.4.1	Ethical Implications of Explainable AI	170
8.4.1.1	Discussion of Ethical Principles	170
8.4.1.2	Fairness	170
8.4.1.3	Accountability	170
8.4.1.4	Patient Privacy	170
8.4.2	Balancing Transparency with Privacy	170
8.4.2.1	Pseudonymization and Data Anonymization	171
8.4.2.2	Granular Access Controls	171
8.4.2.3	Consent and Opt-Out Mechanisms	171
8.4.3	Interdisciplinary Collaboration and User-Centered Design	171
8.4.3.1	Importance of Collaboration	171
8.4.3.2	Ethical Review Boards	171
8.4.3.3	Ethics-Driven Design	172
8.4.4	User-Centered Design Principles	172
8.4.4.1	Human-Centered AI Interfaces	172
8.4.4.2	Incorporating User Feedback	172
8.4.4.3	Empowering User Agency	172
8.4.5	Regular Audits	173

8.4.6	Performance Assessments	173
8.4.7	Feedback Loops	174
8.5	Addressing Ethical Concerns and Algorithmic Biases	175
8.5.1	Ethical Decision-Support Tools: A Framework for Ethical Deliberation	175
8.5.1.1	Ethical Principles and Considerations	175
8.5.1.2	Patient Preferences and Values	176
8.5.1.3	Ethical Dilemma Analysis	176
8.5.1.4	Real-Time Guidance and Feedback	177
8.6	Future Directions and Considerations	177
8.6.1	Summary of Key Points	178
8.6.2	Importance of Explainable AI in Healthcare	178
8.6.3	Benefits of Transparency, Accountability, and Patient Trust	178
8.6.4	Future Directions in Explainable AI Research	178
8.6.5	Interpretability of Complex AI Models	179
8.6.6	Human-Centered Design of Explainable AI Systems	179
8.6.7	Ethical Considerations and Regulatory Compliance	179
8.7	Recommendations for Policymakers, Researchers, and Healthcare Practitioners	179
8.8	Conclusion	180
	References	180
9	Interpretable Machine Learning Techniques	185
	<i>V. Kavitha, K. Suresh, G. Priyadharshini, Shaik Rasheeda Begum and R. Vidhya</i>	
9.1	Introduction	186
9.1.1	Definition of Interpretable Machine Learning	186
9.1.2	Importance of Interpretability in Machine Learning Models	186
9.1.3	Real-World Applications and Implications	187
9.2	Importance of Interpretability	188
9.2.1	Significance in Various Domains	189
9.2.2	Challenges Posed by Complex, Black-Box Models	189
9.2.3	Legal, Ethical, and Regulatory Considerations	190
9.3	Techniques for Interpretable Machine Learning	190
9.3.1	Intrinsic Methods	191
9.3.2	<i>Post-Hoc</i> Methods	191
9.3.3	Hybrid Approaches	191
9.3.3.1	Intrinsic Methods	191

	9.3.3.2	<i>Post-Hoc</i> Methods	192
	9.3.3.3	Hybrid Approaches	192
9.4		Intrinsic Methods	193
	9.4.1	Decision Trees	193
	9.4.1.1	How do Decision Trees Work?	193
	9.4.1.2	Examples of Decision Tree Applications	193
	9.4.2	Linear Models	194
	9.4.2.1	When and How to Use Linear Models for Interpretability	194
	9.4.2.2	Trade-Offs Between Interpretability and Model Performance	194
9.5		<i>Post-Hoc</i> Methods	195
	9.5.1	Illustrative Examples	195
	9.5.2	Partial Dependence Plots	195
	9.5.2.1	Construction and Interpretation	195
	9.5.2.2	Case Studies Demonstrating the Utility of Partial Dependence Plots	196
9.6		Hybrid Approaches	196
	9.6.1	Local Interpretable Model-Agnostic Explanations (LIME)	196
	9.6.1.1	Conceptual Overview	196
	9.6.2	Implementation Details	197
	9.6.3	Practical Examples	197
	9.6.4	SHAP (SHapley Additive exPlanations)	197
	9.6.4.1	How SHAP Values are Computed	197
	9.6.4.2	Use Cases and Benefits	198
9.7		Evaluating Interpretability	198
	9.7.1	Metrics for Evaluating Interpretability	198
	9.7.2	Comparison of Different Techniques	199
	9.7.3	Considerations for Choosing the Appropriate Method	200
9.8		Case Studies: Interpretable Machine Learning in Healthcare	200
	9.8.1	Predicting Patient Readmission Risk	200
	9.8.1.1	Challenges and Lessons Learned	201
	9.8.2	Personalized Treatment Recommendations	201
	9.8.2.1	Overview	201
	9.8.2.2	Challenges and Lessons Learned	201
	9.8.3	Early Detection of Disease Progression	202
	9.8.3.1	Overview of Early Disease Progression	202
	9.8.3.2	Challenges and Lessons Learned	202
9.9		Risk Stratification for Preventive Care	202

9.9.1	Overview of Preventive Care	202
9.9.2	Challenges and Impact	203
9.10	Conclusion	203
	References	204
10	Interpretable Machine Learning Techniques in AI	209
	<i>Shavez, Poornima, Kanu Goyal, Shweta Sharma and Parul Sharma</i>	
10.1	Introduction	210
10.2	History	211
10.3	Making ML Interpretable	212
	10.3.1 Benefits and Challenges to Interpretability	212
	10.3.2 Challenges with Interpretable ML, Including	213
	10.3.3 Limitations of Machine Learning	213
10.4	Machine Learning Models	214
	10.4.1 Local Method Explanation	214
	10.4.2 Global Method Explanation	215
10.5	Techniques of Interpretable Machine Learning	215
10.6	Model of Intrinsic Interpretable	217
	10.6.1 Globally Interpretable Model	217
	10.6.2 Adding Interpretability Constraints	217
	10.6.3 Interpretable Model Extraction	218
	10.6.4 Locally Interpretable Model	218
10.7	<i>Post-Hoc</i> Global Explanation	219
	10.7.1 Traditional ML Explanation	219
	10.7.2 Model Agnostic Explanation	220
	10.7.3 Model-Specific Explanation	220
10.8	DNN Representation Explanation	221
	10.8.1 Explanation of CNN Representation	222
	10.8.2 Explanation of RNN Representation	223
10.9	<i>Post-Hoc</i> Local Explanation	224
	10.9.1 Model-Agnostic Explanation	224
	10.9.2 Explanation Based on Local Approximations	224
	10.9.3 Perturbation-Based Explanation	225
	10.9.4 Model-Specific Explanation	226
10.10	Backpropagation	226
10.11	Mask Perturbation	226
	10.11.1 Investigation of Deep Representations	227
10.12	Conclusion	227
	References	228

11 Interpretable Machine Learning Techniques in Medical System—The Role of Data Analytics and Machine Learning	233
<i>Venkataraman P., Sunantha D. and Lakshmi S.</i>	
11.1 Introduction	234
11.1.1 Classification of Interpretable Methods	236
11.1.2 Disease Diagnosis Using Interpretable Machine Learning	239
11.1.3 Role of Interpretable Machine Learning	239
11.1.3.1 Importance of Interpretability in Healthcare Models	240
11.1.3.2 Challenges in Achieving Interpretability	240
11.1.3.3 Interpretability Techniques	241
11.1.3.4 Human–AI Collaboration	242
11.2 Materials and Methods	243
11.2.1 Diabetes Dataset	243
11.2.2 Machine Learning Algorithms	244
11.2.2.1 Support Vector Machine (SVM)	244
11.2.2.2 KNN	244
11.2.2.3 Random Forest	244
11.2.2.4 Decision Tree	244
11.2.2.5 Naïve Bayes	245
11.2.3 Diabetic Prediction Process	246
11.2.3.1 Data Preprocessing	246
11.2.3.2 Training and Testing	246
11.2.3.3 Implementation of the Classification Model	247
11.3 Experimental Result and Discussion	249
11.3.1 Performance Metrics	249
11.3.1.1 Accuracy	249
11.3.1.2 Confusion Matrix	249
11.3.1.3 F1-Score	250
11.3.1.4 Precision	250
11.3.1.5 Recall	250
11.3.2 Results and Discussion	250
11.4 Conclusion	252
References	252
12 Interpretable AI: Shedding Light on Medical Image Analysis Using Machine Learning Techniques	257
<i>S. Bashyam, P. Supraja and Prithiviraj Rajalingam</i>	
12.1 Introduction	258

12.1.1	Interpretability in AI for Healthcare	258
12.1.2	Scope of the Chapter	258
12.2	Medical Image Analysis: A Critical Application of Machine Learning	259
12.2.1	Significance of Medical Image Analysis	259
12.2.2	Challenges and Complexities in Medical Image Data	259
12.2.3	Role of Machine Learning in Automating Image Analysis	259
12.3	Interpretable AI Techniques	260
12.3.1	Historical Roots of Deep Learning	260
12.3.2	Google Net: A Game-Changer	260
12.3.3	CNN Architecture	260
12.3.4	Types of Learning	260
	12.3.4.1 Learning Problems	261
	12.3.4.2 Hybrid Learning Problems	262
	12.3.4.3 Statistical Inference	263
	12.3.4.4 Learning Techniques	264
12.4	Accelerating Deep Learning Implementation with Open-Source Frameworks	265
12.4.1	Tensor Flow	265
12.4.2	Caffe	265
12.4.3	Caffe 2	266
12.4.4	ONNX (Open Neural Network Exchange)	266
12.4.5	Keras	266
12.4.6	Torch	266
12.4.7	Other Deep Learning Tools and Libraries	266
12.4.8	Additional Libraries and Frameworks	267
12.5	The Machine Learning Landscape in Medical Image Analysis	267
12.5.1	Deep Learning Networks: Tailored Architectures and Objectives	267
12.5.2	Image Registration: Aligning Medical Images for Enhanced Analysis	268
	12.5.2.1 Objective and Techniques	268
	12.5.2.2 Clinical Applications and Perspectives	269
	12.5.2.3 Tools and Toolkits	269
12.5.3	Object Localization in Medical Image Analysis	269
	12.5.3.1 Objective and Techniques	269
	12.5.3.2 Challenges in Localization	270
	12.5.3.3 Novel Strategies	271

12.5.4	Classification and Detection in Medical Image Analysis	271
12.5.4.1	Exam Classification	271
12.5.4.2	Object Classification	271
12.5.4.3	Classification Algorithms	272
12.5.4.4	Algorithm Types Related to Classification	272
12.5.4.5	Important Phrases Used in Classification Algorithms	272
12.5.4.6	Application in Disease Diagnosis	273
12.5.4.7	Transfer Learning for Disease Identification	273
12.5.4.8	Hierarchical Medical Image Classification	273
12.5.5	Object Detection in Medical Image Analysis	274
12.5.6	Segmentation	274
12.5.6.1	Substructure/Organ Segmentation	274
12.5.6.2	Lesion Segmentation	274
12.5.6.3	Role of Medical Image Segmentation	274
12.5.7	Impact of AI and Machine Learning	275
12.6	Use Cases and Applications: Interpretable AI's Role in Brain Tumor Care—Diagnosis, Treatment, and Analysis	275
12.6.1	Precise Diagnosis: Unveiling Clarity	275
12.6.2	Customized Treatment Planning: Excellence in Every Detail	275
12.6.3	Radiomics Analysis: Expanding on It	275
12.6.4	Evaluation of Algorithm Performance	276
12.7	Challenges in Deep Learning for Medical Imaging	277
12.7.1	Imbalanced Data and Lack of Confidence Interval	277
12.7.2	Limited Availability of Annotated Data	277
12.7.3	Transition to End-to-End Learning	277
12.7.4	Data Augmentation Techniques	278
12.7.5	Interpretability and “Black-Box” Models	278
12.7.6	Resource-Intensive Training	278
	References	278
13	Exploring the Role of Explainable AI in Women's Health: Challenges and Solutions	283
	<i>Inam Ul Haq and Akib Mohi Ud Din Khanday</i>	
13.1	Introduction	283

13.2	The Importance of Addressing Women's Health Concerns Beyond Major Diseases	288
13.3	The Role of Explainable AI (XAI) in Providing Precise Medicine Solutions Tailored to Women's Specific Health Needs	289
13.4	Challenges in the Application of Explainable AI in Women's Health	290
13.5	Women's Health Concerns	292
13.6	The Promise of Machine Learning (ML) and Explainable AI (XAI) Technologies in Women's Health	296
13.7	Potential Applications of Machine Learning (ML) and Explainable AI (XAI) Technologies in Women's Health	297
13.8	Case Studies and Examples of Successful Implementations	299
13.9	Specific XAI Techniques and Methodologies	299
13.10	Conclusion	302
	References	303
14	Explainable AI in Healthcare: Introduction	307
	<i>Amandeep Kaur and Sonali Goyal</i>	
14.1	Introduction to AI and Explainable AI	308
14.2	Introduction to Explainable AI in Healthcare	314
14.3	Applications of Explainable AI in Healthcare	317
14.4	Implementing Explainable AI in Healthcare: Practical Considerations	319
14.5	Future Directions and Emerging Trends	319
14.6	Conclusion	320
14.7	Future Work	320
	References	321
15	Ethical Implications of Emotion Recognition Technology in Mental Healthcare: Navigating Privacy, Bias, and Therapeutic Boundaries	325
	<i>R. Ravi, V. Jeya Ramya, B. Prameela Rani, Srikanth Nalluri and M. Jenath</i>	
15.1	Introduction	326
15.2	Background	327
15.2.1	Objectives	328
15.3	Related Works	329
15.4	Literature Survey	330
15.4.1	Capabilities and Applications of ERT	331
15.4.2	Ethical Considerations in ERT	331

15.4.3	Clinical Applications and Efficacy of ERT	331
15.4.4	Challenges and Limitations of ERT	332
15.4.5	Future Directions and Recommendations	332
15.5	Methodology	332
15.5.1	Search Strategy for Literature Review	332
15.5.2	Selection Criteria for Inclusion of Studies	333
15.5.3	Data Extraction Process	334
15.5.4	Analysis of Literature	334
15.5.5	Validation and Peer Review	335
15.5.6	Ethical Considerations	335
15.6	Research Process: Emotion Recognition Technology (ERT) in Mental Healthcare	337
15.6.1	Defining the Research Scope and Objectives	337
15.6.2	Literature Review Protocol Development	337
15.6.3	Search Strategy Formulation	337
15.6.4	Literature Search and Screening	338
15.6.5	Selection of Studies	338
15.6.6	Data Extraction	338
15.6.7	Thematic Analysis	338
15.6.8	Synthesis and Interpretation	338
15.6.9	Validation and Peer Review	339
15.6.10	Ethical Considerations	339
15.7	Research Validation and Methodology Confirmation	339
15.7.1	Research Methodology	339
15.7.2	Research Execution	339
15.7.3	Validation Process	340
15.8	Results	340
15.8.1	Privacy Concerns	341
15.8.2	Algorithmic Biases	341
15.8.3	Accuracy and Validation	341
15.9	Future Directions	343
15.10	Conclusion	345
	References	346
16	Bridging the Gap: Clinical Adoption and User Perspectives of Explainable AI in Healthcare	349
	<i>Shaik Masood Ahamed and J. Jabez</i>	
16.1	Introduction	350
16.2	Background	352
16.3	Research Objectives and Hypotheses	352
16.3.1	Research Objectives	353

	16.3.2 Hypotheses	353
16.4	Related Works	354
16.5	Literature Survey	355
	16.5.1 Technical Foundations of XAI	355
	16.5.2 Impact on Clinical Decision Making	356
	16.5.3 User Perspectives and Acceptance	356
	16.5.4 Challenges and Considerations	356
	16.5.5 Real-World Implementations and Case Studies	356
	16.5.6 Interdisciplinary Perspectives	357
16.6	Proposed Methodology	357
	16.6.1 Study Design	357
	16.6.2 Data Collection Methods	358
	16.6.3 Sampling Strategy	358
	16.6.4 Data Analysis Ways	358
	16.6.5 Ethical Considerations	358
16.7	Project Management	359
	16.7.1 Surveys	359
	16.7.2 Interviews	360
	16.7.3 Observational Studies	360
	16.7.4 Document Analysis	361
	16.7.5 Focus Groups (if Applicable)	361
	16.7.6 Data Analysis Plan	362
16.8	Quantitative Data Analysis	362
	16.8.1 Descriptive Statistics	363
	16.8.2 Inferential Statistics	363
	16.8.3 Regression Analysis	363
	16.8.4 Quantitative Integration	363
	16.8.5 Qualitative Data Analysis	363
	16.8.5.1 Data Coding	363
	16.8.5.2 Theme Development	364
	16.8.5.3 Interpretation and Conflation	364
16.9	Integration of Findings	364
	16.9.1 Case Study: Implementation and Impact of Explainable AI (XAI) in Radiology	364
	16.9.2 Key Research Questions to be Addressed in the Case Study May Include the Following	365
	16.9.3 Literature Survey	366
	16.9.3.1 Technical Foundations of XAI in Radiology	366
	16.9.3.2 Clinical Applications and Impact of XAI in Radiology	366

16.9.3.3	User Perspectives and Acceptance of XAI in Radiology	366
16.9.3.4	Implementation Challenges and Considerations	367
16.9.3.5	Regulatory and Ethical Implications of XAI in Radiology	367
16.9.3.6	Real-World Implementations and Case Studies	367
16.9.3.7	Interdisciplinary Perspectives on XAI in Radiology	367
16.10	Methodology	368
16.10.1	Selection of Research Sites	368
16.10.2	Data Collection Instruments	368
16.11	Results	369
16.11.1	XAI Adoption and Awareness	369
16.11.2	Impact on Diagnostic Accuracy and Efficiency	369
16.11.3	User Perspectives and Experiences	369
16.11.4	Implementation Challenges and Considerations	370
16.11.5	Recommendations for Future Implementation	370
16.12	Conclusion	370
	References	371
17	Application of AI-Based Technologies in the Healthcare Sector: Opportunities, Challenges, and Its Impact—Review	375
	<i>G. Jegadeeswari and B. Kirubadurai</i>	
17.1	Introduction	376
17.2	Literature Review, Methodology, and Analysis	377
17.3	Methodology	382
17.4	Analysis	383
17.5	Discussion and Limitations	387
	Conclusion	389
	References	389
18	A Complete Road Map for Interpretable Machine Learning Techniques Harnessing Various Real-Time Applications	393
	<i>A. Pandian, V. V. Ramalingam, J. Venkata Subramanian, K. Pradeep Mohan Kumar and S. Padmini</i>	
18.1	Introduction	394
18.2	Feature Importance	399
18.3	Rule-Based Models	401
18.4	Model Transparency	403

18.5	Visual Explanation	405
18.6	Interpretable Deep Learning	408
18.7	Evaluating Interpretability	410
18.8	Impact of the Current Study	412
18.9	Conclusion	414
	References	415
19	Future Research Directions: Explainable Artificial Intelligence in Healthcare Industry	423
	<i>Shamneesh Sharma, Neha Kumra, Meghna Luthra, Vikas Verma and Komal Sharma</i>	
19.1	Introduction	424
19.2	Background and Literature Review	426
19.3	Explainable Artificial Intelligence Techniques in Healthcare	429
	19.3.1 Local Interpretable Model-Agnostic Explanations	431
	19.3.2 SHAP (SHapley Additive exPlanations)	432
	19.3.3 Rule-Based XAI Techniques	432
19.4	Applications of XAI in Healthcare	433
19.5	Methodology and Result Analysis	434
	19.5.1 Dataset Information	434
	19.5.2 Research Methodology	435
	19.5.3 Result Analysis	436
19.6	Future Directions	440
	19.6.1 Research Area I Interpretable Deep Learning Models	440
	19.6.2 Research Area II Dynamic and Adaptive Explanations	441
	19.6.3 Research Area III Integration With Electronic Health Records (EHRs)	441
	19.6.4 Research Area IV Human–AI Collaboration	442
	19.6.5 Research Area V: Explainable Genomic Medicine	442
19.7	Conclusion	443
	References	444
20	Real-World Applications of Explainable AI in Healthcare	451
	<i>Urvi, Parul Sharma, Kanu Goyal and Shweta Sharma</i>	
20.1	Introduction	452
	20.1.1 Explainable Artificial Intelligence: Concepts and Techniques	453

20.1.2	Overview of Popular Explainable Artificial Intelligence Techniques	455
20.2	Real-World Applications of Explainable Artificial Intelligence in Healthcare	457
20.2.1	Diagnostic Imaging	457
20.2.2	Clinical Decision Support Systems (CDSS)	458
20.2.3	Drug Discovery and Development	459
20.2.4	Personalized Medicine	460
20.2.5	Patient Outcome Prediction	460
20.3	Challenges and Ethical Considerations for Explainable Artificial Intelligence Usage in Healthcare	461
20.4	Future Scope	462
20.5	Conclusion	463
	References	463
21	Explainable AI in Medical Imaging, Personalized Medicine, and Bias Reduction: A New Era in Healthcare	467
	<i>Komal, Ganesh K. Sethi, Shamneesh Sharma and Rajender Kumar</i>	
21.1	Introduction	468
21.2	Explainable AI Techniques	470
21.3	Challenges in Healthcare AI	474
21.4	Complexity of Machine Learning Models	476
21.5	Explainable AI Techniques	477
21.6	Benefits of XAI in Healthcare	479
21.7	Improved Trust and Adoption by Healthcare Professionals	481
21.8	Case Studies of Explainable Artificial Intelligence	483
21.9	Challenges and Limitations	486
21.10	Future Directions	487
21.11	Conclusion	488
	References	489
22	Understanding Explainability in Medical Imaging	493
	<i>Annie Silviya S. H., R. Tamizh Kuzhali, Akshaya V., Lakshmi Prabha T. S., Immanuvel Arokia James K. and B. Sriman</i>	
22.1	Introduction	494
22.1.1	Definition of Explainability in Medical Imaging	494
22.1.2	Importance of Explainability in Healthcare	494
22.1.3	Overview of the Chapter Structure	496

22.2	The Role of Medical Imaging in Healthcare	497
22.2.1	Brief History and Evolution of Medical Imaging Technologies	497
22.2.2	Current Significance of Medical Imaging in Healthcare	499
22.3	Explainability Challenges in Medical Imaging	500
22.3.1	Complexity of Medical Imaging Data and Techniques	500
22.3.2	Black-Box Nature of Deep Learning Models	501
22.3.3	Ethical and Regulatory Considerations in Medical AI	503
22.4	Techniques for Explainability in Medical Imaging	504
22.4.1	Feature Importance Methods	504
22.4.2	Saliency Maps and Gradient-Based Methods	504
22.4.3	Model-Agnostic Techniques	505
22.4.4	Rule-Based Systems	506
22.4.4.1	Case Study 1: Melanoma Classification with SHAP	506
22.4.4.2	Clinical Implications	507
22.4.4.3	Clinical Implications	508
22.5	Interpretability vs. Explainability	508
22.5.1	Importance of Providing Understandable Explanations	509
22.5.2	Advantages and Limitations of Explainability in Medical Imaging	510
22.5.3	Advantages of Explainable AI in Medical Imaging	510
22.5.3.1	Enhancing Trust and Transparency	510
22.5.3.2	Facilitating Collaboration Between AI Systems and Healthcare Professionals	510
22.5.4	Limitations and Challenges of Explainable AI in Medical Imaging	511
22.5.4.1	Trade-Offs Between Explainability and Performance	511
22.5.4.2	Interpreting Complex Models	511
22.5.4.3	Generalization and Robustness	511
22.6	Regulatory Landscape and Standards	512
22.6.1	Regulatory Frameworks in AI Healthcare	512
22.6.2	Standardization Efforts in AI Healthcare	513
22.6.3	Challenges and Future Directions	513
22.6.4	Future Directions	514

22.7	Case Studies and Applications of Explainable AI in Medical Imaging	515
22.7.1	Success Stories and Lessons Learned	516
	Conclusion	517
	References	517
23	Explainability and Regulatory Compliance in Healthcare: Bridging the Gap for Ethical XAI Implementation	521
	<i>Uma Maheswari Kalia Moorthy, Asthampatti Marimuthu Jayapalan Muthukumaran, Vijayalakshmi Kaliyaperumal, Shobana Jayakumar and Kalpana Ayanellore Vijayaraghavan</i>	
23.1	Introduction	522
23.2	Explainability Techniques and Methodologies	527
23.2.1	Real-World Examples of Explainable AI in Healthcare	529
23.3	Regulatory Landscape in Healthcare	531
23.3.1	Healthcare Regulations and Policies Related to AI Implementation	533
23.3.2	Data Privacy and Patient Confidentiality Considerations	535
23.3.3	Legal and Ethical Challenges in Deploying AI in Healthcare	537
23.3.3.1	Privacy and Security	537
23.3.3.2	Informed Consent	537
23.3.3.3	Bias and Fairness	538
23.3.3.4	Transparency and Explainability	538
23.3.3.5	Liability and Accountability	538
23.3.3.6	Clinical Validation and Regulation	538
23.3.3.7	Human Oversight and Autonomy	539
23.3.3.8	Consent for Research and Innovation	539
23.3.3.9	Equity and Access	539
23.4	Ensuring Regulatory Compliance	540
23.5	Strategies to Comply with Healthcare Regulations and Standards	540
23.6	The Role of Interdisciplinary Collaboration in Regulatory Compliance	542
23.7	Addressing Ethical Concerns and Mitigating Bias in AI Models	543
23.8	Trust and Transparency in Healthcare AI	546

23.8.1	Building Trust with Patients and Healthcare Professionals Through Explainable AI	547
23.8.2	The Impact of Transparency on AI Adoption and User Acceptance	547
23.8.3	Human–AI Interaction and the Role of Explanations in Decision Making	548
23.8.4	AI Projects in Healthcare That Lacked Transparency or Compliance	548
23.8.5	Ethical Considerations in Explainable AI Implementation	550
23.8.6	Ethical Principles for Explainable AI in Healthcare	551
23.8.7	Ensuring Patient Safety and Well-Being Through Explainable AI	553
23.8.8	Ethical Challenges in Balancing Privacy, Transparency, and Utility in Healthcare AI	555
23.9	Future Trends and Recommendations	557
	References	558
24	Envisioning Explainable AI: Significance, Real-Time Applications, and Challenges in Healthcare	563
	<i>Kannan Chakrapani, Mohamed Iqbal Safa, Saranya Gangadharra Moorthy, Meenakshi Kumaraswamy and George Parimala</i>	
24.1	Introduction	564
24.2	Definition and Significance of XAI	565
24.3	Significance of Explainable AI (XAI)	565
24.4	Interpretable Decision Support Systems	574
24.5	Applications of XAI in Healthcare Industry	576
24.5.1	Application in Medical Diagnosis	578
24.6	Applications of XAI in Finance	579
24.7	Applications of XAI in Judiciary	582
24.8	Applications in the Judiciary	583
24.9	Challenges and Limitations of XAI	585
24.10	Limitations of XAI	585
24.11	Future Directions in XAI	586
	Conclusion	588
	References	588

25 Enlightened XAI: Illuminating Ethics and Equitable Explainability	593
<i>Hemalatha P., Manikandan J., B. Balaji and V. Sujitha</i>	
25.1 Introduction	594
25.1.1 Definition of XAI	596
25.2 Ethical Considerations of AI in Healthcare	598
25.3 The Need for Ethical Guidelines in AI in Healthcare Development	599
25.4 Risks Associated with Opaque and Black-Box AI in Healthcare Models	601
25.5 Transparency and Interpretability in XAI	602
25.5.1 Importance of Transparency in Decision-Making Algorithms	602
25.5.2 Interpretable Models vs. Complex Models	602
25.5.3 Techniques for Achieving Interpretability in AI	603
25.6 Fairness in AI	604
25.6.1 Understanding Algorithmic Bias	604
25.7 Consequences of Biased AI Decision Making	604
25.8 Real-World Examples of AI Fairness Issues	606
25.9 Fairness in Explainable AI	608
25.10 The Tradeoff Between Fairness and Interpretability	609
25.11 Evaluating Fairness in Explainable AI Models	610
25.12 Guidelines and Frameworks for Ethical XAI	611
25.12.1 Existing Principles and Guidelines for Responsible XAI Development	611
25.12.2 Regulatory Considerations and Legal Frameworks	611
25.12.3 Industry Initiatives Promoting Ethical XAI	612
25.13 The Future of Ethical XAI	612
25.14 Conclusion	614
References	614
26 Enhancing Trust and Collaboration Using Explainability in Natural Language Processing for AI-Driven Healthcare	619
<i>A. Pandian, K. Pradeep Mohankumar, S. Padmini, Sibi Amaran and K. Sreekumar</i>	
26.1 Introduction	620
26.1.1 Background of NLP in Healthcare	621
26.2 Importance of Explainability in NLP	622
26.3 Challenges of Black-Box NLP Models in Healthcare	625
26.4 Opacity of NLP Models	627
26.5 Addressing the Incompatibility of NLP Models	629

26.6	Implications of Lack of Explainability in Medical Decision Making	629
26.7	Necessity for Explainable AI in Healthcare	631
26.8	Building Trust and Confidence in NLP Predictions	632
26.9	Validation and Accountability in Medical Applications	633
26.10	Techniques for Achieving Explainability in NLP	634
26.10.1	Rule-Based Explanations	635
26.10.2	Visualizations	636
26.10.3	Grad-CAM (Gradient-weighted Class Activation Mapping)	637
26.10.4	Textual Explanation Generation	638
26.11	Advantages of Explainability for Healthcare Professionals	640
26.11.1	Empowering Healthcare Practitioners With AI Insights	640
26.11.2	Enhancing Collaboration between AI and Medical Experts	643
26.12	Future Directions and Challenges	645
26.13	Challenges Associated With Explainable AI for Healthcare	645
	Conclusion	646
	References	647
	About the Editors	651
	Index	653

