

**ACADEMIA DE STUDII ECONOMICE DIN BUCUREȘTI**

**Consiliul pentru Studii Universitare de Doctorat**

**Școala Doctorală de Finanțe**

**Machine Learning aplicat în Finanțe**

Codruț-Florin Ivașcu

Conducător de doctorat:  
prof. univ. dr. Bogdan Negrea

**București, 2024**

## Rezumat

Această teză își propune să examineze rolul și aplicabilitatea algoritmilor de Machine Learning în domeniul financiar, oferind o evaluare echilibrată între promisiunile tehnologiei și limitările acesteia. Lucrarea navighează între entuziasmul teoretic și realismul aplicativ, încercând să construiască o punte între capacitatea tehnologică și nevoile practice ale sectorului financiar. Motivația principală a lucrării este de a depăși metodele tradiționale prin demonstrarea modului în care ML poate oferi soluții superioare în numeroase domenii financiare.

Într-o perioadă definită de accelerarea inovațiilor tehnologice, Inteligența Artificială și ML au capturat atenția publicului și a cercetătorilor, fiind adesea prezentate ca soluții omnipotente pentru provocările economice și sociale. În acest context, această teză argumentează că o înțelegere profundă a principiilor de bază, a potențialului și a restricțiilor ML este esențială pentru aplicarea sa eficientă în finanțe. În această direcție, lucrarea demarează cu o discuție detaliată a diferitelor modele de ML, de la modele liniare și metode de regularizare, la algoritmi avansați bazați pe arbori, rețele neuronale, și metode de ansamblu. Discuția este amplificată de analiza avantajelor și limitărilor fiecărui model, oferind o perspectivă asupra potențialului lor de a transforma analiza financiară și gestionarea deciziilor.

Aplicabilitatea ML este ulterior ilustrată prin multiple studii de caz care acoperă diferite aspecte ale finanțelor, demonstrând capacitatea tehnologiei de a depăși metodele tradiționale. Studiile de caz includ evaluarea opțiunilor, unde algoritmi precum XGBoost și LightGBM oferă evaluări mai precise decât modelele parametrice consacrate; gestionarea portofoliilor, unde tehnici de ML optimizează selecția și alocarea activelor; managementul riscului de credit, care beneficiază de o evaluare a riscului mai precisă; modelarea politicii de dividend, unde ML poate identifica factorii determinanți care influențează politicile companiilor într-o manieră nonliniară; modelarea economiei subterane, care utilizează ML pentru a înțelege interacțiunile complexe dintre diferitele cheltuieli guvernamentale; și prognoza inflației, care prezintă eficacitatea modelelor ML în predicția inflației comparativ cu metodele econometrice tradiționale.

Concluziile tezei reflectă asupra dualității ML: pe de o parte, un instrument potent pentru inovare și eficiență în sectorul financiar; pe de altă parte, un domeniu plin de provocări tehnice care necesită o gestionare prudentă. Teza subliniază necesitatea unei abordări critice și informate în aplicarea ML, accentuând că succesul acesteia depinde nu doar de algoritmi utilizați, ci și de integritatea datelor, precizia modelării și relevanța contextului aplicativ.

# CUPRINS

<b>SECȚIUNEA I – INTRODUCERE .....</b>	<b>1</b>
<b>1. CONSIDERENTE GENERALE .....</b>	<b>2</b>
1.1. MOTIVAȚIE .....	2
1.2. CE ESTE ML? .....	5
1.3. STRUCTURA TEZEI.....	7
<b>2. ISTORIE.....</b>	<b>12</b>
2.1. VIZIUNI PREMODERNE.....	12
2.2. ÎNCEPUTURILE AI .....	13
2.3. SISTEMELE EXPERT .....	14
2.4. REVOLUȚIA MACHINE LEARNING.....	16
2.5. BIG DATA ȘI ASCENSIUNEA DEEP LEARNING .....	17
2.6. IMPACTUL SOCIAL ȘI ETIC .....	18
<b>3. CONCEPTE FUNDAMENTALE .....</b>	<b>20</b>
3.1. TIPURI DE ML .....	20
3.2. PROCESUL DE MACHINE LEARNING.....	22
3.3. EVALUAREA ȘI OPTIMIZAREA MODELELOR .....	23
3.4. INTERPRETABILITATEA MODELELOR.....	27
3.4.1. SHAP .....	28
3.4.2. LIME.....	29
3.4.3. ALTE METODE.....	31
<b>SECȚIUNEA II – MODELE .....</b>	<b>33</b>
<b>4. MODELE LINIARE .....</b>	<b>34</b>
4.1. INTRODUCERE .....	34

<b>4.2. REGRESIA LINIARA .....</b>	<b>35</b>
4.2.1. FORMULAREA MATEMATICĂ .....	35
4.2.2. PROPRIETĂȚILE ESTIMATORILOR.....	36
4.2.3. IPOTEZE ALE REGRESIEI LINIARE.....	37
4.2.4. AVANTAJE ȘI DEZAVANTAJE .....	38
<b>4.3. REGRESIA LOGISTICĂ .....</b>	<b>39</b>
4.3.1. FORMULAREA MATEMATICĂ .....	40
4.3.2. REGRESIA LOGISTICĂ MULTINOMIALĂ .....	41
4.3.3. IPOTEZELE MODELULUI.....	41
4.3.4. AVANTAJE ȘI DEZAVANTAJE .....	42
<b>4.4. METODE DE REGULARIZARE .....</b>	<b>43</b>
4.4.1. LASSO.....	43
4.4.2. RIDGE .....	46
4.4.3. ELASTIC NET .....	47
<b>4.5. CONCLUZII .....</b>	<b>48</b>
<b><u>5. SUPPORT VECTOR MACHINES.....</u></b>	<b><u>49</u></b>
<b>5.1. INTRODUCERE .....</b>	<b>49</b>
<b>5.2. ISTORIE.....</b>	<b>49</b>
<b>5.3. PREZENTAREA MODELULUI .....</b>	<b>51</b>
5.3.1. FORMULAREA MATEMATICĂ .....	52
5.3.2. TEHNICI KERNEL ÎN SVM.....	54
5.3.3. REGULARIZARE ȘI OPTIMIZARE .....	55
<b>5.4. SUPPORT VECTOR REGRESSION (SVR).....</b>	<b>56</b>
5.4.1. FORMULAREA MATEMATICĂ .....	57
5.4.2. REGULARIZARE ȘI OPTIMIZARE.....	58
<b>5.5. AVANTAJE ȘI DEZAVANTAJE.....</b>	<b>59</b>
<b>5.6. CONCLUZII .....</b>	<b>60</b>
<b><u>6. ARBORI DE DECIZIE .....</u></b>	<b><u>62</u></b>
<b>6.1. DECISION TREE (MODELUL CART).....</b>	<b>62</b>
6.1.1. INTRODUCERE.....	62
6.1.2. ISTORIE.....	63
6.1.3. PREZENTAREA MODELULUI.....	65
6.1.4. ASPECTE TEHNICE ALE ALGORITMULUI.....	70

6.1.5. AVANTAJE SI DEZAVANTAJE.....	71
<b>6.2. RANDOM FOREST .....</b>	<b>72</b>
6.2.1. INTRODUCERE.....	72
6.2.2. ISTORIE.....	73
6.2.3. PREZENTAREA MODELULUI.....	74
6.2.4. ASPECTE TEHNICE ALE ALGORITMULUI.....	77
6.2.5. AVANTAJE ȘI DEZAVANTAJE.....	78
6.2.6. CONCLUZII.....	79
<b>6.3. XGBOOST.....</b>	<b>80</b>
6.3.1. INTRODUCERE.....	80
6.3.2. ISTORIE.....	81
6.3.3. PREZENTAREA MODELULUI.....	82
6.3.4. CARACTERISTICILE DISTINCTIVE ALE XGBOOST.....	86
6.3.5. HIPERPARAMETRII .....	90
6.3.6. AVANTAJE ȘI DEZAVANTAJE.....	91
6.3.7. CONCLUZII.....	92
<b>6.4. LIGHTGBM.....</b>	<b>93</b>
6.4.1. INTRODUCERE.....	93
6.4.2. CARACTERISTICI DISTINCTIVE .....	94
6.4.3. AVANTAJE ȘI DEZAVANTAJE .....	98
6.4.4. CONCLUZII.....	99
<b><u>7. REȚELE NEURONALE.....</u></b>	<b><u>100</u></b>
<b>7.1. INTRODUCERE .....</b>	<b>100</b>
<b>7.2. ISTORIE.....</b>	<b>101</b>
<b>7.3. NEURONUL ARTIFICIAL .....</b>	<b>104</b>
<b>7.4. DEEP NEURAL NETWORK – MULTILAYER PERCEPTRON.....</b>	<b>106</b>
<b>7.5. ANTRENAREA .....</b>	<b>108</b>
7.5.1. FUNCȚIA OBIECTIV.....	108
7.5.2. BACKPROPAGATION .....	109
7.5.3. ALGORITMI DE OPTIMIZARE.....	110
7.5.4. ÎNȚIALIZAREA PONDERILOR.....	113
7.5.5. REGULARIZARE .....	113
<b>7.6. AVANTAJE ȘI DEZAVANTAJE.....</b>	<b>115</b>
<b>7.7. TIPURI AVANSATE DE REȚELE.....</b>	<b>117</b>
7.7.1. REȚELE NEURONALE RECURENTE (RNN) .....	117

7.7.2. LONG SHORT TERM MEMORY (LSTM).....	119
<b>7.8. CONCLUZII .....</b>	<b>122</b>
<b><u>8. ALTE MODELE SUPERVIZATE .....</u></b>	<b><u>124</u></b>
<b>8.1. K-NEAREST NEIGHBOURS (KNN) .....</b>	<b>124</b>
8.1.1. INTRODUCERE.....	124
8.1.2. DESCRIEREA MODELULUI.....	125
8.1.3. AVANTAJE ȘI DEZAVANTAJE.....	126
<b>8.2. ALGORITMI GENETICI.....</b>	<b>127</b>
8.2.1. INTRODUCERE.....	127
8.2.2. DESCRIEREA MODELULUI.....	128
8.2.3. AVANTAJE ȘI DEZAVANTAJE.....	134
<b><u>9. MODELE NESUPERVIZATE .....</u></b>	<b><u>136</u></b>
<b>9.1. INTRODUCERE .....</b>	<b>136</b>
<b>9.2. REDUCEREA DIMENSIONALITĂȚII.....</b>	<b>137</b>
9.2.1. PRINCIPAL COMPONENT ANALYSIS (PCA).....	138
9.2.2. AUTOENCODERI.....	140
9.2.3. ALTE MODELE.....	142
<b>9.3. CLUSTERING.....</b>	<b>143</b>
9.3.1. K-MEANS .....	145
9.3.2. HIERARCHICAL CLUSTERING .....	146
9.3.3. DBSCAN .....	147
9.3.4. ALTE MODELE.....	148
<b><u>SECȚIUNEA III – STUDII DE CAZ .....</u></b>	<b><u>150</u></b>
<b><u>10. STUDIU DE CAZ I – EVALUAREA OPTIUNILOR.....</u></b>	<b><u>151</u></b>
<b>10.1. INTRODUCERE .....</b>	<b>151</b>
<b>10.2. METODOLOGIE.....</b>	<b>154</b>
10.2.1. DATE.....	154
10.2.2. BENCHMARKS.....	155
10.2.3. CONSIDERENTE GENERALE .....	158
<b>10.3. REZULTATE.....</b>	<b>162</b>

<b>10.4. CONCLUZII .....</b>	<b>167</b>
<b><u>11. STUDIU DE CAZ II – MANAGEMENTUL PORTOFOLIULUI.....</u></b>	<b><u>170</u></b>
<b>11.1. INTRODUCERE .....</b>	<b>170</b>
<b>11.2. SELECȚIA.....</b>	<b>171</b>
11.2.1. RANDOM FOREST.....	172
11.2.2. XGBOOST.....	173
11.2.3. LASSO.....	173
11.2.4. AUTOENCODERI.....	174
<b>11.3. ALOCAREA.....</b>	<b>175</b>
11.3.1. OLS.....	176
11.3.2. SENZITIVITATEA REȚELELOR NEURONALE.....	176
<b>11.4. METODOLOGIE.....</b>	<b>177</b>
<b>11.5. REZULTATE.....</b>	<b>179</b>
<b>11.6. CONCLUZII.....</b>	<b>186</b>
<b>11.7. ANEXĂ.....</b>	<b>187</b>
<b><u>12. STUDIU DE CAZ III – MANAGEMENTUL RISCULUI.....</u></b>	<b><u>193</u></b>
<b>12.1. INTRODUCERE .....</b>	<b>193</b>
<b>12.2. STADIUL CUNOAȘTERII .....</b>	<b>194</b>
<b>12.3. METODOLOGIE.....</b>	<b>197</b>
12.3.1. DATE.....	197
12.3.2. TEHNICI DE REBALANSARE .....	199
12.3.3. CONSIDERENTE GENERALE .....	200
12.3.4. METRICI DE EVALUARE .....	201
<b>12.4. REZULTATE.....</b>	<b>202</b>
<b>12.5. CONCLUZII .....</b>	<b>209</b>
<b><u>13. STUDIU DE CAZ IV – MODELAREA POLITICII DE DIVIDEND .....</u></b>	<b><u>210</u></b>
<b>13.1. INTRODUCERE .....</b>	<b>210</b>
<b>13.2. STADIUL CUNOAȘTERII .....</b>	<b>211</b>
<b>13.3. METODOLOGIE.....</b>	<b>215</b>
13.3.1. ANALIZA DATELOR .....	215
13.3.2. CONSIDERENTE GENERALE .....	219

13.4. REZULTATE .....	221
13.5. CONCLUZII .....	225
<b><u>14. STUDIU DE CAZ V – MODELAREA ECONOMIEI SUBTERANE .....</u></b>	<b><u>226</u></b>
14.1. INTRODUCERE .....	226
14.2. STADIUL CUNOAȘTERII .....	227
14.3. DESCRIEREA DATELOR.....	230
14.4. ANALIZA LINIARĂ .....	233
14.5. ANALIZA NELINIARĂ.....	235
14.6. CONCLUZII .....	241
<b><u>15. STUDIU DE CAZ VI – PREDICTIA INFLATIEI.....</u></b>	<b><u>243</u></b>
15.1. INTRODUCERE .....	243
15.2. METODOLOGIE.....	244
15.3. REZULTATE.....	245
15.4. CONCLUZII .....	248
<b><u>SECȚIUNEA IV – DISCUȚII .....</u></b>	<b><u>249</u></b>
<b><u>16. CONCLUZII GENERALE.....</u></b>	<b><u>250</u></b>
<b><u>17. ESTE ML UN PANACEU UNIVERSAL? .....</u></b>	<b><u>257</u></b>
<b><u>REFERINȚE BIBLIOGRAFICE .....</u></b>	<b><u>262</u></b>
<b><u>LISTA TABELELOR .....</u></b>	<b><u>280</u></b>
<b><u>LISTA FIGURILOR.....</u></b>	<b><u>282</u></b>
<b><u>GLOSAR DE TERMENI.....</u></b>	<b><u>284</u></b>