

بسمه تعالی



دانشگاه آزاد اسلامی

این قسمت توسط حوزه معاونت

پژوهشی دانشگاه پر می شود.

شماره:

تاریخ:

پیوست:

### فرم طرح تحقیق

دکترای حرفه‌ای  کارشناسی ارشد

درخواست تصویب موضوع پایان‌نامه کارشناسی ارشد و دکترای حرفه‌ای

توجه: این فرم با مساعدت و هدایت استاد راهنما تکمیل شود.

عنوان تحقیق به فارسی:

انتخاب هوشمند ویژگیها در الگوریتم های برنامه نویسی ژنتیک جهت تشخیص بیماری سرطان سینه

عنوان تحقیق به انگلیسی:

**Intelligent selection features in the genetic programming algorithm to detect breast cancer**

۱. اطلاعات مربوط به دانشجو

نام خانوادگی:	نام:
شماره دانشجویی:	رشته تحصیلی: کامپیوتر
گرایش: نرم افزار	مقطع: کارشناسی ارشد
دانشکده: تحصیلات تکمیلی	دوره: ارشد ناپیوسته
نیمسال و سال ورود:	نشانی پستی منزل:
	نشانی پستی محل کار:
تلفن همراه:	تلفن ثابت:

۲. اطلاعات مربوط به استاد راهنما

نام خانوادگی:	نام:
آخرین مدرک تحصیلی:	تخصص جنبی:
رتبه و پایه دانشگاهی:	سنوات تدریس: کارشناسی ارشد: ..... دکتری: .....
تلفن همراه:	نحوه همکاری: تمام وقت <input type="checkbox"/> نیمه وقت <input type="checkbox"/> مدعو <input type="checkbox"/>
تلفن:	نشانی:

تعداد پایان‌نامه‌های کارشناسی ارشد راهنمایی شده:

سایر دانشگاه‌ها:

دانشگاه آزاد اسلامی:

نام پایان‌نامه‌های کارشناسی ارشد راهنمایی در یک سال گذشته:

دانشگاه آزاد اسلامی:

سایر دانشگاه‌ها:

تعداد پایان‌نامه‌های کارشناسی ارشد در دست راهنمایی:

سایر دانشگاه‌ها:

دانشگاه آزاد اسلامی:

نام پایان‌نامه‌های کارشناسی ارشد در دست راهنمایی:

دانشگاه آزاد اسلامی:

سایر دانشگاه‌ها:

تعداد رساله‌های راهنمایی شده در یک سال گذشته:

سایر دانشگاه‌ها:

دانشگاه آزاد اسلامی:

نام رساله‌های راهنمایی شده دکترا:

دانشگاه آزاد اسلامی:

سایر دانشگاه‌ها:

تعداد رساله‌های دکترا در دست راهنمایی:

سایر دانشگاه‌ها:

دانشگاه آزاد اسلامی:

نام رساله‌های دکترا در دست راهنمایی:

دانشگاه آزاد اسلامی:

سایر دانشگاه‌ها:

### ۳. اطلاعات مربوط به استادان مشاور

نام :	نام خانوادگی :	تخصص اصلی :
آخرین مدرک تحصیلی :	رتبه و پایه دانشگاهی :	تلفن :
شغل :		
محل خدمت :		
تعداد پایان‌نامه‌ها و رساله‌های راهنمایی شده : کارشناسی ارشد : ..... دکترا : .....		
تعداد پایان‌نامه‌ها و رساله‌های در دست راهنمایی : کارشناسی ارشد : ..... دکترا : .....		

نام :	نام خانوادگی :	تخصص اصلی :
آخرین مدرک تحصیلی :	رتبه و پایه دانشگاهی :	تلفن :
شغل :		
محل خدمت :		
تعداد پایان‌نامه‌ها و رساله‌های راهنمایی شده : کارشناسی ارشد : ..... دکترا : .....		
تعداد پایان‌نامه‌ها و رساله‌های در دست راهنمایی : کارشناسی ارشد : ..... دکترا : .....		

### ۴. اطلاعات مربوط به پایان‌نامه

الف : عنوان پایان‌نامه : انتخاب هوشمند ویژگی‌ها در الگوریتم‌های برنامه نویسی ژنتیک جهت تشخیص بیماری سرطان سینه

فارسی  غیرفارسی

ب: نوع کار تحقیقاتی: بنیادی<sup>۱</sup>  نظری<sup>۲</sup>  کاربردی<sup>۳</sup>  عملی<sup>۴</sup>

ت: پرسش اصلی تحقیق (مسئله تحقیق) : چگونه می‌توان به کمک برنامه نویسی ژنتیک ویژگی‌های مهم و کلیدی را در تشخیص بیماری سرطان سینه بکار گرفت؟

۱- تحقیق بنیادی: پژوهشی است که به کشف ماهیت اشیاء، پدیده‌ها و روابط بین متغیرها، اصول، قوانین و ساخت با آزمایش‌های تئوریک و نظریه‌ها می‌پردازد و به توسعه مرزهای دانش رشته علمی کمک می‌نماید.

۲- تحقیق نظری: نوعی پژوهش بنیادی است و از روشهای استدلال و تحلیل عقلانی استفاده می‌کند و بر پایه مطالعات کتابخانه‌ای انجام می‌شود.

۳- تحقیق کاربردی: پژوهشی است که با استفاده از نتایج تحقیقات بنیادی به منظور بهبود و به کمال رساندن رفتارها، روشها، ابزارها، وسایل، تولیدات، ساختار و الگوهای مورد استفاده جوامع انسانی انجام می‌شود.

۴- تحقیق علمی: پژوهشی است که با استفاده از نتایج تحقیقات بنیادی و با هدف رفع مسائل و مشکلات انسانی انجام می‌شود.

۵. بیان مسأله ( تشریح ابعاد، حدود مسأله، معرفی دقیق مسأله، بیان جنبه‌های مجهول و مبهم و متغیرهای و متغیرهای مربوط به پرسشهای تحقیق، منظور از تحقیق )

سرطان سینه<sup>۱</sup> در حال حاضر یکی از شایع‌ترین بیماری‌های کشنده کشورهای در حال توسعه است که جان نیمی از افراد جامعه را تهدید می‌نماید. شیوع سرطان سینه در زنان به نسبت تقریبی یک به پانصد است که بیشترین مرگ و میر در سنین بین ۳۵ تا ۶۵ سال اتفاق می‌افتد. سرطان سینه در مراحل اولیه کشنده نمی‌باشد و به مرور زمان به وسیله جریان خون، بافت‌های سرطانی<sup>۲</sup> در سایر اندام‌های حیاتی بدن به پخش شده و منجر به فراگیری سلول‌های سرطانی در بدن و نهایتاً مرگ بیمار می‌باشد. عوامل مختلفی نظیر وراثت<sup>۳</sup>، مصرف الکل، چاقی، سطح بالای استروژن<sup>۴</sup>، بارداری در سنین بالا، در معرض بودن انواع اشعه ایکس<sup>۵</sup> و فرابنفش<sup>۶</sup> و غیره در بروز این سرطان کشنده موثر می‌باشند. تشخیص سرطان سینه با انجام یک نمونه‌برداری از توده مربوطه تأیید می‌شود. پس از تشخیص سرطان سینه، آزمایش‌های بیشتری انجام می‌شوند تا مشخص شود که آیا سرطان به قسمت‌های دیگر بدن نیز سرایت کرده است یا خیر و چه درمان‌هایی ممکن است نسبت به بیماری واکنش نشان دهند. آزمایشات منظم غربالگری<sup>۷</sup> که به کمک ماموگرافی<sup>۸</sup> صورت می‌پذیرد، می‌تواند جهت تشخیص زود هنگام بیماری و درمان استفاده شود. اعمال برنامه غربالگری<sup>۹</sup> جهت تشخیص سرطان سینه باعث کاهش مرگ میر زنان می‌شود لذا در بسیاری از کشورهای توسعه‌یافته برنامه غربالگری به شکل منظم و گسترده‌ای اعمال می‌شود. متأسفانه به دلیل مختلفی نظیر عدم آموزش بانوان، هزینه‌های غربالگری سینه، ترس از آزمایشات بالینی و غیره استقبال بالایی از این روش تشخیص سرطان سینه نمی‌شود. در بروز سرطان سینه عوامل مختلفی وجود دارد که تحت نام ویژگی<sup>۱۰</sup> بیماری شناخته می‌شوند و مسلماً تعدادی مشخص از این ویژگی‌ها نقش مهم و کلیدی در تشخیص بیماری سرطان سینه را ایفاء می‌نمایند. انتخاب درست و دقیق ویژگی‌ها باعث می‌شود که تشخیص بیماری سرطان سینه از اطلاعات بیماران امکان پذیر شود و با کاوش در میان داده‌های موجود در مورد بیماری بتوان به الگوهای درست این بیماری دست پیدا کنیم. داده‌کاوی<sup>۱۱</sup> یک روش و تکنیک موثر برای کشف الگوهای پنهان<sup>۱۲</sup> در داده‌های غیر سازماندهی شده و حجیم نظیر بیماری سرطان سینه می‌باشد. داده‌کاوی در ترکیب با الگوریتم‌های هوشمند تکاملی نظیر برنامه‌نویسی ژنتیک<sup>۱۳</sup> توانایی مناسبی برای تشخیص ویژگی و الگوهای بیماری دارد لذا در این پژوهش قصد داریم به کمک برنامه‌نویسی ژنتیک ویژگی‌های مهم این بیماری را انتخاب نموده و تاثیر آنها را در پیش‌بینی بیماری سرطان سینه را مورد بررسی قرار دهیم.

1 Breast cancer

2 Cancer tissues

3 Inheritance

4 Estrogen

5 X Ray

6 UV radiation

7 Breast screening

8 Mammography

9 Breast screening programme

10 Features

11 Data minig

12 Hidden patterns

13

۶. سوابق مربوط (بیان مختصر سابقه تحقیقات انجام شده درباره موضوع و نتایج به دست آمده در داخل و خارج از کشور نظرهای علمی موجود درباره موضوع تحقیق)

شیخی پور و همکاران (۲۰۱۶) در [۱]، با استفاده از الگوریتم بهینه‌سازی ذرات<sup>۱۴</sup> و کرنل تخمین چگالی<sup>۱۵</sup> یک روش جدید در انتخاب ویژگی‌های مهم مرتبط با بیماری سرطان سینه ارائه نمودند. نشان دادند که انتخاب ویژگی به کمک ترکیب الگوریتم ذرات و کرنل تخمین چگالی ویژگی‌ها موثرتر از انتخاب ویژگی به کمک ترکیب الگوریتم ژنتیک و کرنل تخمین چگالی می‌باشد. نتایج بررسی آنها نشان داد انتخاب ویژگی به کمک الگوریتم ذرات در تشخیص سرطان سینه دقتی در حدود ۹۸.۴۵٪ را ارائه می‌دهد و این در حالی است که دقت الگوریتم ژنتیک نیز ۹۸.۴۵٪ است اما با تعداد ویژگی‌های بیشتری که بکار می‌رود.

ساسیکالا<sup>۱۶</sup> و همکاران (۲۰۱۵) در [۲]، جهت تشخیص سرطان سینه از مجموعه ژن‌های که می‌تواند در ارتباط با این سرطان جهش یابد یک تکنیک مبتنی بر برنامه‌نویسی ژنتیک را ارائه نمودند. در روش پیشنهادی آنها هر کروموزم زیر مجموعه‌ای از ژن‌ها در نظر گرفته می‌شود و به عنوان بردار ویژگی سرطان سینه مورد استفاده قرار می‌گیرد. در روش پیشنهادی آنها اگر دو کروموزم که حاوی ویژگی‌های مسئله می‌باشد، دارای شایستگی یکسانی باشند آنگاه کروموزمی که تعداد ویژگی کمتری را در خود دارد به عنوان کروموزم شایسته‌تر انتخاب می‌شود. روش پیشنهادی آنها در واقع یک جستجوی هوشمندانه جهت تشخیص الگوی ژن‌های فعال در بیماری سرطان سینه می‌باشد. آنها برای تشخیص الگوی سرطان سینه از ترکیب نسخه پیشنهادی الگوریتم ژنتیک با ابزارهای داده‌کاوی نظیر درخت تصمیم‌گیری، ماشین بردار پشتیبان، شبکه بیزی و نزدیکترین همسایه استفاده نمودند که نتایج پژوهش آنها نشان می‌دهد که دقت در این روش‌ها به ترتیب ۹۳.۸۱٪، ۹۱.۷۵٪، ۸۸.۵٪ و ۸۲.۴۷٪ می‌باشد. همچنین نتایج بررسی و پژوهش آنها نشان می‌دهد که تعداد ویژگی از ۲۴۴۸۱ به ۶ ویژگی کلیدی در تشخیص بیماری تقلیل یافت.

میلونکوویچ<sup>۱۷</sup> و همکاران (۲۰۱۵) در [۳]، یک روش ناحیه‌بندی جهت شناسایی پستان و جداسازی بافتهای آن از قفسه سینه و بافتهای مجاور به کمک فیلتر گابور<sup>۱۸</sup> ارائه نمودند. آنها به کمک کوتاهترین مسیر<sup>۱۹</sup> موجود در یک نواحی و ارزش‌گذاری با یک تابع هزینه جدید توانستند با دقتی در حدود ۹۶.۱٪ ساختار پستان را از بدن جداسازی نمایند. در روش پیشنهادی آنها، در ابتدا یک تصویر به کمک الگوریتم خوشه‌بندی FCM به تعدادی ناحیه از پیش تعیین شده تقسیم می‌شود و به موازات این عمل لبه‌یابی به کمک فیلتر گابور صورت می‌گیرد. به کمک خروجی الگوریتم خوشه‌بندی فازی مجموعه‌ای از نقاط نشانه<sup>۲۰</sup> برای جداسازی بافت چربی سینه از بدن انتخاب می‌شود و خروجی این مجموعه نقاط در اختیار الگوریتم کوتاه‌ترین مسیر قرار می‌گیرد. یکی از مزایای

<sup>14</sup> particle swarm optimization

<sup>15</sup> kernel density estimation (KDE)

<sup>16</sup> Sasikala

<sup>17</sup> Jana Milenković

<sup>18</sup> Gabor filter

<sup>19</sup> shortest-path search

<sup>20</sup> Seed point

مهم روش پیشنهادی فیلتر گابور و کوتاه‌ترین مسیر در تشخیص بافت سینه، عدم نیاز به هر گونه داده آموزشی است که می‌تواند روند سرعت اجرای برنامه را کاهش دهد.

هانگ<sup>۲۱</sup> و همکاران (۲۰۱۵) در [۴]، یک روش جدید جهت استخراج بافت‌های توده‌ای سینه به کمک تصاویر فرا صوت پستان<sup>۲۲</sup> ارائه نمودند. روش پیشنهادی آنها شامل چند مرحله اساسی است. در مرحله اول آنها به کمک مدل رفع نویز بر اساس اختلاف کل شدت نور تصویر<sup>۲۳</sup> مقدار نویز تصویر را کاهش دادند و مزیت استفاده از این روش علاوه بر کاهش نویز، بهبود کیفیت لبه‌های ناحیه تومور می‌باشد. در مرحله دوم یک ناحیه-بندی مبتنی بر گراف صورت گرفته و تصویر به زیر نواحی مختلفی تقسیم می‌شود که توسط کانال RGB رنگی و متمایز می‌شود. در ادامه با توجه به تصویر ناحیه‌بندی شده مبتنی بر گراف ویژگی‌های عمده تصویر استخراج شده و در نهایت کاهش ویژگی اتفاق افتاد و بر اساس بردار ویژگی کاهش یافته و در نهایت مرحله طبقه‌بندی نواحی مختلف پستان انجام می‌شود. روش پیشنهادی آنها با ۲۵ ویژگی بر روی مجموعه‌ای از تصاویر بکار گرفته شد که نتایج نشان می‌دهد دقت، حساسیت و تشخیص آن برای استخراج تومور پستان در حدود ۹۸.۳٪، ۹۷٪ و ۹۸.۵٪ می‌باشد.

ساناتان‌مان<sup>۲۴</sup> و همکاران (۲۰۱۵) در [۵]، برای تشخیص بیماری دیابت از یک الگوریتم ترکیبی خوشه‌بندی Kmeans به کمک الگوریتم ژنتیک<sup>۲۵</sup> استفاده نمودند. در روش پیشنهادی آنها در ابتدا مقادیر پوچ<sup>۲۶</sup> و خالی به کمک مقادیر میانگین ویژگی‌های مرتبط به روزرسانی می‌شود و در ادامه به کمک الگوریتم خوشه‌بندی ابعاد مسئله کاهش داده می‌شود و همزمان به کمک الگوریتم تکاملی ژنتیک تعداد ویژگی‌های مهم بیماری دیابت استخراج شده و از این ویژگی‌ها در ماشین بردار پشتیبان<sup>۲۷</sup> استفاده می‌شود تا افراد به دو دسته سالم و بیمار طبقه‌بندی شوند. نتایج بررسی‌ها نشان می‌دهد دقت روش آنها بالا (مزیت روش) و در حدود ۹۸.۷۹٪ درصد است و این درحالی است که روش ترکیبی Kmeans و ماشین بردار پشتیبان دارای دقتی در حدود ۹۶.۷۱٪ می‌باشد. یکی از معایب این روش ترکیب سه الگوریتم زمان‌بر و زمان اجرای بالای الگوریتم تشخیصی آنها می‌باشد.

وآنگ<sup>۲۸</sup> و همکاران (۲۰۱۵) در [۶]، یک روش انتخاب ویژگی<sup>۲۹</sup> در تشخیص بیماری دیابت به وسیله الگوریتم آزمون صدای مخالف<sup>۳۰</sup>، الگوریتم اکتشافی الکترومغناطیس<sup>۳۱</sup> و الگوریتم نزدیکترین همسایه<sup>۳۲</sup> را ارائه نمودند. روش پیشنهادی آنها برای انتخاب ویژگی با نه روش انتخاب ویژگی بیماری دیابت مورد مقایسه قرار گرفته شد که نتایج آزمایشات نشان می‌دهد انتخاب ویژگی به روش مورد نظر می‌تواند دقت بیماری دیابت

<sup>21</sup> Qinghua Huang

<sup>22</sup> Breast ultrasound (BUS)

<sup>23</sup> total-variation (TV) restoration model

<sup>24</sup> Santhanam

<sup>25</sup> Genetic algorithm (GA)

<sup>26</sup> Null

<sup>27</sup> Support Vector Machine (SVM)

<sup>28</sup> Wang

<sup>29</sup> feature selection

<sup>30</sup> opposite sign test (OST)

<sup>31</sup> Electromagnetism-like mechanism

<sup>32</sup> Nearest neighbor algorithm

نوع دو<sup>۳۳</sup> را افزایش دهد. مزیت عمده روش آنها بررسی مجموعه داده‌های مختلف برای شناسایی ویژگی‌های مهم در بیماری دیابت و عیب روش آنها عدم در نظر گرفتن هر ویژگی و وزن آن در جوامع مختلف با سبک زندگی متفاوت می‌باشد.

جابهز<sup>۳۴</sup> و همکاران (۲۰۱۵) در [۷]، جهت تشخیص انواع بیماری‌ها مانند سرطان کبد، دیابت، بیماری قلبی، سرطان سینه و دیابت مجموعه‌ای از ابزارهای داده‌کاوی را با الگوریتم تکاملی ذرات ترکیب شده با الگوریتم ژنتیک<sup>۳۵</sup> را بکار نمودند. در روش پیشنهادی آنها الگوریتم تکاملی نقش افزایش میزان دقت طبقه‌بندی افراد سالم و بیمار را به عهده دارد و در واقع ابزارهای داده‌کاوی را نظیر درخت تصمیم‌گیری دقیق‌تر می‌نماید. به عنوان نمونه نتایج آزمایشات نشان داد که در مورد بیماری کبد از ۴۳ قانون با دقت ۶۲.۸۹٪ در روش‌های مبتنی بر درخت تصمیم‌گیری به ۱۹ قانون با کیفیت‌تر با دقت ۶۴.۶۰٪ در روش پیشنهادی حاصل شد. در مورد بیماری قلبی از ۹۸ قانون با دقت ۶۸.۶۴٪ در روش‌های مبتنی بر درخت تصمیم‌گیری به ۱۹ قانون با کیفیت‌تر با دقت ۶۴.۶۰٪ در روش پیشنهادی حاصل شد. در مورد بیماری قلبی از ۹۸ قانون با دقت ۶۸.۶۴٪ در روش پیشنهادی حاصل شد. در مورد بیماری قلبی از ۹۸ قانون با دقت ۶۴.۶۰٪ در روش پیشنهادی حاصل شد. نتایج بررسی‌های آنها در مورد بیماری دیابت و هیپاتیت نیز امیدوارکننده و با دقت مناسبی ارائه شده است. مزیت اصلی این مقاله بررسی تعدادی بیماری با یک روش پیشنهادی و مقایسه روش پیشنهادی در این بیماری‌ها و مقایسه جامع آنها می‌باشد و عیب روش آنها عدم استفاده از سایر تکنیک‌های مهم داده‌کاوی نظیر شبکه عصبی مصنوعی در تشخیص بیماری‌ها می‌باشد.

موهاپاترآ و همکاران (۲۰۱۵) در [۸]، به کمک تکنیک یادگیری شدید<sup>۳۶</sup> که یکی از روش‌های مهم تشخیص موهاپاترآ در بیماری‌ها است و الگوریتم بهبود یافته فاخته<sup>۳۷</sup>، یک روش جدید جهت تشخیص بیماری‌های مانند دیابت، سرطان سینه و هیپاتیت ارائه و توسعه دادند. در روش پیشنهادی آنها وزن و آستانه شبکه یادگیرنده به وسیله الگوریتم بهبود یافته فاخته به گونه‌ای انتخاب می‌شود که میزان خطای طبقه‌بندی به عنوان تابع هدف کمینه ممکن شود. نتایج پژوهش آنها نشان می‌دهد که روش ترکیبی آنها فادر است با ۹۷.۷٪ سرطان سینه، ۷۸.۵٪ بیماری دیابت و ۱۰۰٪ بیماری دیابت را تشخیص دهد. مزیت عمده روش آنها ترکیب الگوریتم تکاملی فاخته با چند الگوریتم یادگیری ماشین و مقایسه ۱۰ تکنیک یادگیری ماشین در تشخیص بیماری‌های نظیر سرطان سینه، بیماری دیابت و هیپاتیت می‌باشد. از جمله معایب روش پیشنهادی آنها عدم انتخاب ویژگی بیماری‌ها بصورت هوشمندانه است.

ویلیکالا<sup>۳۸</sup> و همکاران (۲۰۱۵) در [۹]، یک روش انتخاب ویژگی هوشمندانه با استفاده از الگوریتم ژنتیک و ماشین بردار پشتیبان در جهت تشخیص بیماری اختلال بینایی دیابت<sup>۳۹</sup> که رشد غیر طبیعی عروق خونی در شبکیه چشم را به علت بیماری دیابت در پی دارد، ارائه و توسعه دادند. در روش پیشنهادی آنها بردارهای

<sup>33</sup> Type 2 diabetes mellitus (DM)

<sup>34</sup> Jabez

<sup>35</sup> Wind-driven Swarm Optimization (WSO)

<sup>36</sup> Extreme learning machine (ELM)

<sup>37</sup> improved cuckoo search (ICS)

<sup>38</sup> Welikala

<sup>39</sup> Proliferative diabetic retinopathy (PDR)

ویژگی اولیه بیماری اختلال دیابت از روی مجموعه‌ای از تصاویر انتخاب شده و در اختیار الگوریتم ماشین بردار پشتیبان قرار داده می‌شود و در هر تکرار الگوریتم مقدار شایستگی بردار ویژگی محاسبه می‌شود و در ادامه به کمک الگوریتم ژنتیک بردارهای ویژگی به روزرسانی شده و مجدداً به کمک تکنیک ماشین بردار پشتیبان برای طبقه‌بندی تصاویر بکار گرفته می‌شود و مراحل بالا تا تکرار آخر الگوریتم ژنتیک ادامه دارد و در نهایت بهترین کروموزم که بهترین بردار ویژگی است برای طبقه‌بندی تصاویر بکار گرفته می‌شود. نتایج ارزیابی آنها نشان می‌دهد که حساسیت و تشخیص روش بر روی ۶۰ تصویر مختلف به ترتیب ۹۱.۳۸٪ و ۹۶٪ می‌باشد که مقادیری بسیار نزدیک به ۱۰۰٪ می‌باشند (مزیت). یکی از معایب روش‌ها آنها تعداد تصاویر بار رفته در روش پیشنهادی آنها است که تعداد آنها اندک است و از طرفی این روش نیاز به زمان نسبتاً بالایی برای پردازش تصاویر و یادگیری ماشین دارد.

علوی و همکاران (۲۰۱۵) در [۱۰]، با استفاده از ترکیب شبکه عصبی مصنوعی با نسخه‌های مختلف الگوریتم کرم شب‌تاب نظیر کرم شب‌تاب ترکیب شده با الگوریتم گرم و سرد<sup>۴۰</sup>، الگوریتم کرم شب‌تاب پروازی<sup>۴۱</sup> و الگوریتم کرم شب‌تاب پروازی به همراه فرآیند گرم و سرد<sup>۴۲</sup> (روش پیشنهادی آنها) یک روش تشخیص الگو در مجموعه داده‌های مختلف نظیر دیابت و بیماری قلبی ارایه نمودند. نتایج آزمایشات نشان می‌دهد که روش پیشنهادی آنها دارای دقتی در حدود ۹۰.۱۰٪ در تشخیص بیماری دیابت، ۸۲.۳۵٪ در بیماری قلبی و ۹۳.۸۸٪ در بیماری پارکینسون<sup>۴۳</sup> می‌باشد. از مزیت‌های عمده روش پیشنهادی آنها انتخاب ویژگی مهم شبکه عصبی به کمک الگوریتم کرم شب‌تاب و دقت مناسب این روش در تشخیص دقیق بیماری‌ها است. زمان آموزش و مشخص نبودن تعداد بهینه لایه‌های پنهان در شبکه عصبی از معایب روش آنها محسوب می‌شود زیرا هیچ پیشنهادی برای انتخاب تعداد نورون‌های پنهان در این پژوهش نشده است.

#### ۷. فرضیه‌ها یا پرسشهای تحقیق

- ۱- برنامه نویسی ژنتیک قادر است که ویژگی‌های مهم بیماری سرطان سینه را با دقت بالا انتخاب نماید.
- ۲- افزایش تعداد تکرار و اندازه جمعیت در الگوریتم ژنتیک شانس یافتن ویژگی‌های با تعداد کم و در حین حالا با تاثیر زیاد بر روی بیماری را افزایش می‌دهد.
- ۳- روش پیشنهادی دارای دقت، تشخیص و حساسیت مناسبی در تشخیص بیماری سرطان سینه است.

#### ۸. اهداف تحقیق ( شامل اهداف علمی ، کاربردی ، و ضرورت‌های خاص انجام تحقیق )

تشخیص بیماری سرطان سینه با روش‌های داده‌کاوی و مفاهیم تکاملی

<sup>40</sup> firefly algorithm with simulated annealing (denoted as SFA)

<sup>41</sup> Lévy flight within the firefly algorithm (denoted as LFA)

<sup>42</sup> integrating SFA with Lévy flight (denoted as LSFA)

<sup>43</sup> Parkinsons



۹. در صورت داشتن هدف کاربردی بیان نام بهره‌وران (اعم از مؤسسات آموزشی و اجرایی و غیره) مراکز درمانی، پژوهشکده‌ها، مراکز تحقیقاتی مرتبط با بیماری‌ها و غیره

۱۰. جنبه نوآوری و جدید بودن تحقیق در چیست؟ ( این قسمت توسط استاد راهنما تکمیل شود.)

در این پژوهش با یک روش جدید مهمترین ویژگی‌های مرتبط با بیماری سرطان سینه به کمک برنامه‌نویسی ژنتیک استخراج می‌شود و از این ویژگی‌ها برای تشخیص دقیق بیماری سرطان سینه استفاده می‌شود.

امضاء استاد

۱۱. روش کار

الف: نوع روش تحقیق: نوع روش تحقیق کتابخانه‌ای است

ب. روش گردآوری اطلاعات (میدانی، کتابخانه‌ای و غیره): روش گردآوری اطلاعات نیز کتابخانه‌ای است.

پ. ابزار گردآوری اطلاعات (پرسشنامه، مصاحبه، مشاهده، آزمون، فیش، جدول، نمونه‌برداری، تجهیزات آزمایشگاهی و بانکهای اطلاعاتی و شبکه‌های کامپیوتری و ماهواره‌ای و غیره): کتب مرجع، پایان‌نامه‌های مرتبط، مقالات داخلی و خارجی

ث. روش تجزیه و تحلیل اطلاعات: تجزیه و تحلیل بکار رفته در این تحقیق به کمک شبیه‌سازی روش پیشنهادی و نتایج حاصل از آن خواهد بود.

۱۲. جدول زمان‌بندی مراحل انجام دادن تحقیق از زمان تصویب تا دفاع نهایی

تا تاریخ	از تاریخ	تاریخ تصویب
		مطالعات کتابخانه‌ای
		جمع‌آوری اطلاعات
		تجزیه و تحلیل داده‌ها
		نتیجه‌گیری و نگارش پایان‌نامه
		تاریخ دفاع نهایی
		طول مدت اجرای تحقیق: ۶ ماه

### ۱۳. مشخصات منابع و مأخذ مورد استفاده در پایان نامه به ترتیب زیر:

در مورد کتاب : نام خانوادگی نویسنده، نام نویسنده ، سال نشر، عنوان کتاب، مترجم، محل انتشار، جلد

در مورد مقاله : نام خانوادگی، نام، عنوان مقاله، عنوان نشریه، سال، دوره، شماره، صفحه

1. Sheikhpour, R., Sarram, M. A., & Sheikhpour, R. (2016). Particle swarm optimization for bandwidth determination and feature selection of kernel density estimation based classifiers in diagnosis of breast cancer. *Applied Soft Computing*, 40, 113-131.
2. Sasikala, S., alias Balamurugan, S. A., & Geetha, S. (2015). A Novel Feature Selection Technique for Improved Survivability Diagnosis of Breast Cancer. *Procedia Computer Science*, 50, 16-23.
3. Milenković, J., Chambers, O., Mušič, M. M., & Tasič, J. F. (2015). Automated breast-region segmentation in the axial breast MR images. *Computers in biology and medicine*, 62, 55-64.
4. Huang, Q., Yang, F., Liu, L., & Li, X. (2015). Automatic segmentation of breast lesions for interaction in ultrasonic computer-aided diagnosis. *Information Sciences*, 314, 293-310.
5. Santhanam, T., & Padmavathi, M. S. (2015). Application of K-Means and Genetic Algorithms for Dimension Reduction by Integrating SVM for Diabetes Diagnosis. *Procedia Computer Science*, 47, 76-83.
6. Wang, K. J., Adrian, A. M., Chen, K. H., & Wang, K. M. (2015). An improved electromagnetism-like mechanism algorithm and its application to the prediction of diabetes mellitus. *Journal of biomedical informatics*, 54, 220-229.
7. Christopher, J. J., Nehemiah, H. K., & Kannan, A. (2015). A Swarm Optimization approach for clinical knowledge mining. *Computer methods and programs in biomedicine*, 121(3), 137-148.
8. Mohapatra, P., Chakravarty, S., & Dash, P. K. (2015). An improved cuckoo search based extreme learning machine for medical data classification. *Swarm and Evolutionary Computation*, 24, 25-49.
9. Welikala, R. A., Fraz, M. M., Dehmeshki, J., Hoppe, A., Tah, V., Mann, S., ... & Barman, S. A. (2015). Genetic algorithm based feature selection combined with dual classification for the automated detection of proliferative diabetic retinopathy. *Computerized Medical Imaging and Graphics*, 43, 64-77.
10. Alweshah, M., & Abdullah, S. (2015). Hybridizing firefly algorithms with a probabilistic neural network for solving classification problems. *Applied Soft Computing*, 35, 513-524.