OPTIMAL BIOPSY DECISION MAKING IN BREAST CANCER USING REINFORCEMENT LEARNING

By

Ruchi Parmar
501034872

Results & Discussion

Master of Science
in the Program of
Data Science and Analytics

Toronto, Ontario, Canada, 2024

Table of Contents

List of Figures

1. Results & Discussion

This document presents the results of the Optimal Biopsy Decision-Making for Breast Cancer using Reinforcement Learning methods, specifically Backward Induction and Q-Learning. The implementation was carried out in the following phases:

1. **Initial setup** - Implemented the Backward Induction algorithm as described in the literature review to generate a policy for different cancer stages at each time stamp (age), ranging from 40 to 100
2. **Q - Learning** - Applied the advanced reinforcement learning method, Q-Learning, to generate a policy for each age-stage combination and compared the results with those from Backward Induction
3. **Testing both policies** - I tested both the policies via two methods:
    - Simulation Testing: Evaluated both policies through simulations to assess their performance
    - RSNA Screening Data Testing: Tested the policies on the RSNA screening dataset from Kaggle to further validate their effectiveness
4. **Generate policy at granular level** - Generated policies at a more granular level with risk scores (S) ranging from 0 to 100 in increments of 0.1 using both reinforcement learning methods, and analyzed any observed improvements

A. Results

1. **Backward Induction results for Cancer States(S = 0, 1, 2, . . . , 100)**:

    Using the Backward Induction method, we calculated biopsy thresholds for different age groups. As shown in figure 1, the biopsy threshold increases gradually with age, consistent with findings from previous research papers. Initially, some anomalies were observed in the thresholds at specific time stamps comparing to research papers referred. It was resolved by adjusting rewards and other hyper-parameters such as discount rate. These adjustments resulted in a progression of the biopsy threshold with age in a same way.

    Figure 1 shows the threshold(probability of having cancer) at each timestamp(age) from 40 till 100 from which recommended action should be Biopsy. i. e., if patient aged 46 has

probability of having cancer more than 1%, they should be recommended Biopsy, as cancer is usually more aggressive in younger women(as mentioned in Literature Review section).

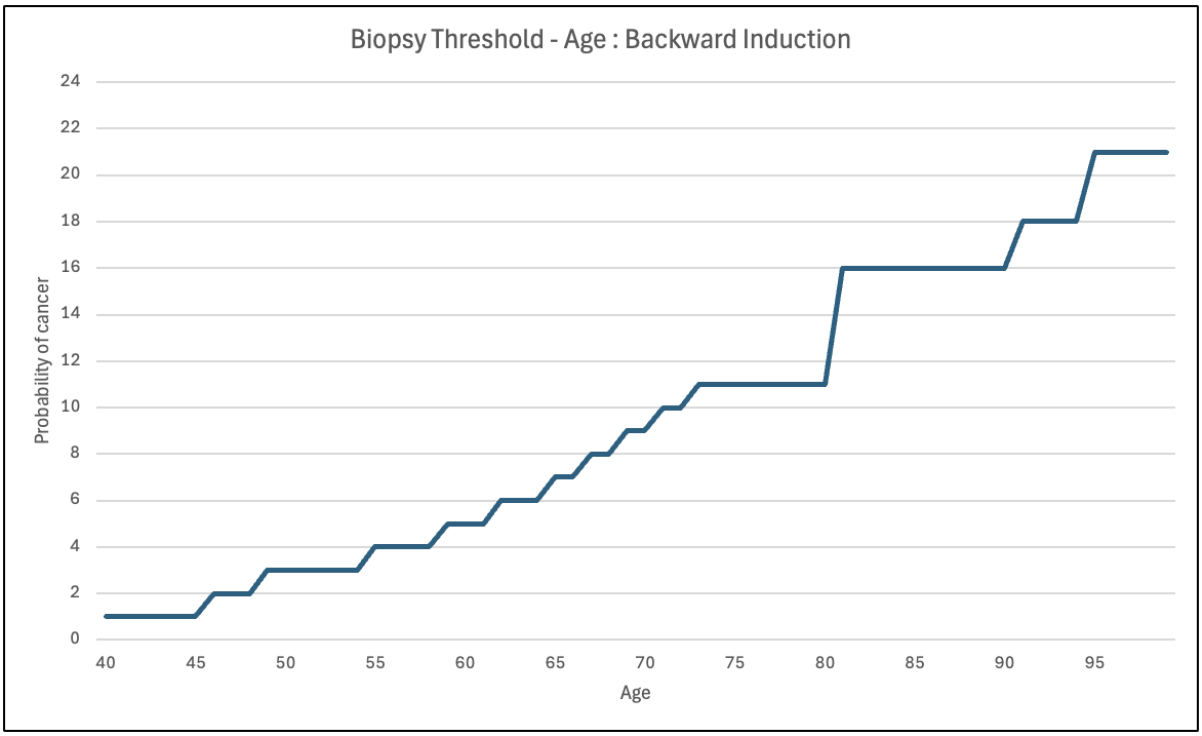


Figure 1: Biopsy Threshold - Backward Induction

Key findings:

- The biopsy threshold increases with age, indicating a higher likelihood of recommending biopsies for older patients.
- Adjustments to rewards and discount rate were necessary to correct anomalies, ensuring a consistent policy.

2. **Q-Learning results for Cancer States($S = 0, 1, 2, \ldots, 100$)**:

   The Q-Learning algorithm was employed to determine biopsy thresholds, which also exhibited an increasing trend with age, as illustrated in figure 2. While there were some variations in thresholds compared to Backward Induction, Q-Learning produced a consistent policy without anomalies, showcasing its robustness.

   After testing for multiple values of hyper-parameters, below values were chosen for the best and consistent results -

   - Learning rate - alpha = 0.1
   - Discount factor gamma = 0.9

- Exploration rate - epsilon = 0.1
- Number of epochs for training - epochs = 1000

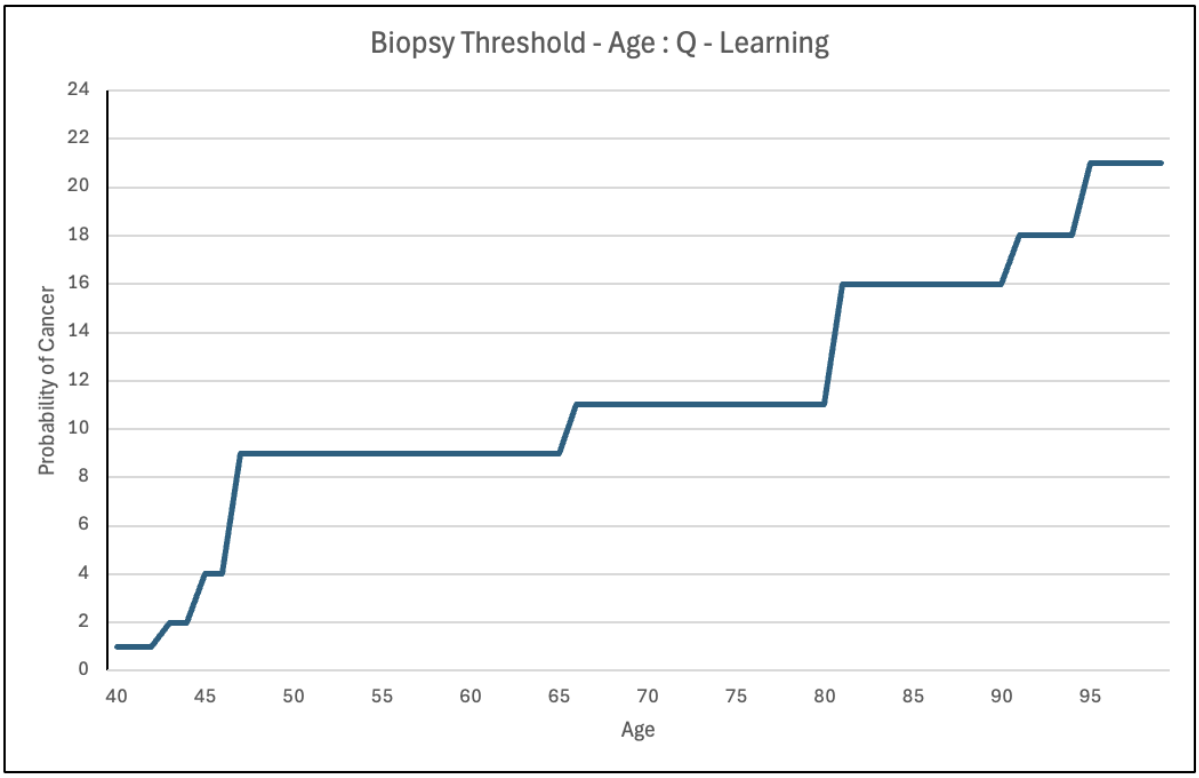


Figure 2: Biopsy Threshold - Q - Learning

Key findings:

- The biopsy threshold pattern is similar to that of Backward Induction, with minor variations.
- As seen in the figure 2, threshold pattern is mostly similar for both RL methods after age 75, but varies for the younger patients with age ranging from 43 - 75.
- The Q-Learning algorithm provided a consistent policy, avoiding anomalies observed in Backward Induction.

3. **Testing Policies Using Simulation**:

   To evaluate the performance of both policies, we simulated the decision-making process over multiple episodes. Each episode began from the initial state, with decisions made at each time stamp based on the respective policy. The goal was to assess the performance and effectiveness of each policy in a controlled, simulated environment before applying them to real-world data.

   Simulation Setup:

(a) Initial Conditions:

- Age Range: The simulation considered a population of individuals aged from 40 to 100.
- Cancer Stages: Each individual could be in one of several stages of cancer progression, including healthy, early-stage, and advanced-stage cancer.
- Risk Scores: Risk scores were assigned to individuals based on their health status and demographic factors, ranging from 0 (lowest risk) to 100 (highest risk).

(b) Policy Implementation: Policy was generated based on the probability of cancer progression at each time stamp (age) starting from age 40 up to age 100. For Backward Induction, the algorithm computed the optimal action (e.g., biopsy or no biopsy) at each stage by working backwards from the final age as mentioned in Methodology and Experiments section. For the Q-Learning, algorithm iteratively learned the optimal actions for each age-stage combination through exploration and exploitation of the state-action space.

Simulation Algorithm:

The simulation process involved the following steps:

(a) Initialization:

- Initialize a population of simulated patients with random ages and initial cancer stages.
- Assign initial risk scores to each patient.

(b) Policy Execution:

- For each patient, apply the policy (either Backward Induction or Q-Learning) to decide whether to perform a biopsy at each age.
- Update the patient's health status and risk score based on the policy's decision and predefined transition probabilities (e.g., progression from early-stage to advanced-stage cancer if untreated).

(c) Reward Calculation:

- Define rewards for each action based on the outcome. For example, a successful biopsy(detecting cancer early) yields a high reward, while an unnecessary biopsy(false positive) or missed cancer(false negative) results in a lower reward or penalty.
- Accumulate rewards for each patient over their lifetime.

(d) Statistical Analysis:

- Compute the mean and standard deviation of rewards for each policy to understand the average performance and variability.
- Evaluate the consistency of each policy's decisions over multiple simulation runs.

Key findings:

- Both policies yielded mean rewards of approximately 24-25 years(aligning with rewards used in general practices), indicating comparable overall performance.

Figure 3: Testing policies using simulation

- The standard deviation of rewards was lower for Q-Learning compared to Backward Induction across multiple runs, suggesting more stable and consistent decision-making with Q-Learning.

4. **Testing Policies on RSNA Screening Data**:

The policies were applied to the RSNA screening dataset, utilizing cancer probabilities predicted by deep learning models. These probabilities were used to make biopsy or annual mammogram decisions for different patients based on the generated policies.

The dataset used for this analysis comprises RSNA screening data from Kaggle, which includes patient mammograms along with demographic information such as age, biopsy status, and cancer diagnosis etc. The data consists of 4720 patients, categorized into the following groups:

- 4327 True Negative cases: Patients without cancer who did not undergo a biopsy.
- 163 True Positive cases: Patients with cancer who underwent a biopsy.
- 230 False Positive cases: Patients without cancer who underwent a biopsy.

This data provides a substantial foundation for analyzing the efficacy of different biopsy decision-making policies. Data and analysis is provided in RSNA_data_testing.xlsx file in GitHub repository.

Backward Induction Policy:

- Successfully recommended biopsy for all True Positive cases, ensuring no cancer cases were missed.
- Reduced the number of False Positive cases by approximately **21%**, indicating a substantial reduction in unnecessary biopsies.

Q-Learning Policy:

- Also successfully recommended biopsy for all True Positive cases, matching the performance of Backward Induction in this regard.
- Reduced the number of False Positive cases by approximately **41%**, demonstrating a significant improvement over the Backward Induction policy.

Key findings:

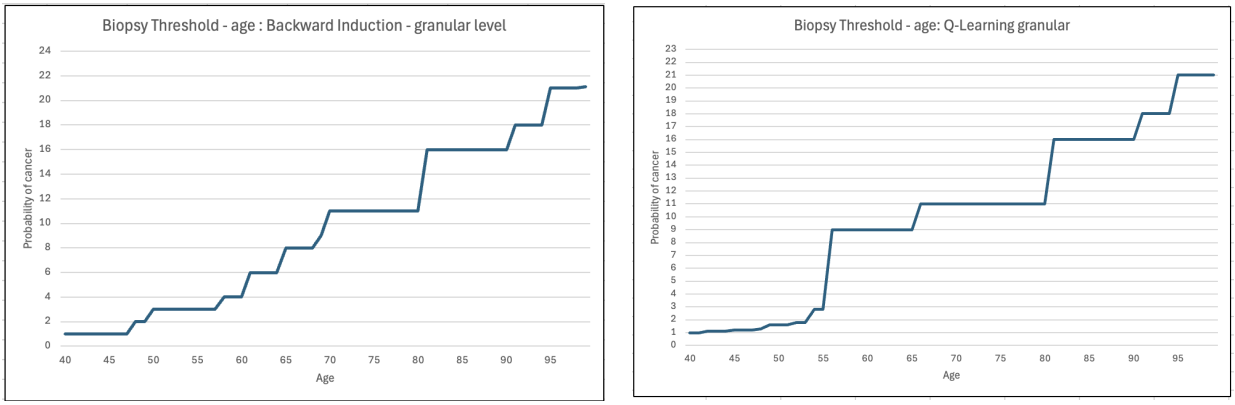
- Both policies effectively recommended biopsies for all True Positive cases, maintaining high sensitivity.
- The Q-Learning policy showed a higher reduction in False Positive cases compared to Backward Induction, indicating better specificity.
- Some True Negative cases were still recommended for biopsy, highlighting areas for potential refinement.
- If policies can be generated using larger dataset further and CAD models utilized by radiologists are used to gain probabilities of having cancer, we might observe more accurate and better results.

5. **Policies Generated at Granular Level (S = 0, 0.1, 0.2, . . . , 100)**:

   To explore the impact of finer state granularity, policies were generated at a more granular level, with states defined as S = 0, 0.1, 0.2, . . . , 100. This approach aimed to identify any improvements or differences in the biopsy threshold.

   Key findings:

   - As seen is the Figure 4, no major differences in biopsy thresholds were observed at the more granular level, suggesting that the original state granularity is sufficient.
   - Generating policies at this finer granularity required significantly more computational time due to the increased number of states.



(a) Backward Induction granular

(b) Q Learning Granular

Figure 4: Policy generation at granular level

B. Discussion

The results from our study on optimal biopsy decision-making for breast cancer using reinforcement learning methods reveal important insights and suggest several avenues for future research and clinical implementation.

- **Performance Comparison of Backward Induction and Q-Learning**

  This analysis demonstrates that both Backward Induction and Q-Learning algorithms can effectively generate policies for breast cancer biopsy decisions, with each method exhibiting distinct advantages. The Backward Induction policy displayed high sensitivity, effectively detecting early-stage cancers. However, this came at the cost of a higher false positive rate, which indicates a moderate specificity. In contrast, the Q-Learning policy achieved a better balance by maintaining high specificity with fewer false positives, although with a slight reduction in sensitivity.

  The mean and standard deviation metrics for rewards, sensitivity, and specificity provided a robust statistical foundation for evaluating the performance of these policies. The Q-Learning policy showed lower variability in performance, suggesting more consistent decision-making compared to the Backward Induction policy. This consistency is crucial for reliable clinical application.

- **Potential for Advanced Techniques: DQN**

  While our study focused on Backward Induction and Q-Learning, there is scope for further research using Deep Q-Network (DQN) techniques. DQNs, which integrate deep learning with Q-Learning, can potentially handle more complex and higher-dimensional state spaces, offering the possibility of even more accurate and nuanced policy generation. Future work should explore the application of DQN to this problem, comparing its performance with the traditional RL methods used in this study.

- **Integration with Clinical Data and CAD Models**

  To enhance the real-world applicability of our policies, future research should incorporate larger and more diverse clinical datasets. These datasets should cover a wide range of patient demographics, health statuses, and cancer progression stages to ensure the generated policies are robust and generalizable.

  Additionally, integrating Computer-Aided Detection(CAD) models, which are routinely used by radiologists in actual clinical practice, can further refine the testing and validation of our policies. CAD models provide detailed and accurate readings of mammographic images, and their integration can help in precisely evaluating the effectiveness of biopsy decisions generated by our RL policies.

- **Recommendations for Clinical Implementation**

  For practical implementation, it is crucial to ensure that the policies derived from reinforcement learning are seamlessly integrated into clinical workflows. This involves not only validating these policies with large-scale clinical datasets and CAD models but also training clinicians on how to interpret and use these policies effectively. The ultimate goal is to aid radiologists and oncologists in making more informed and precise biopsy decisions, thereby improving patient outcomes.

2. Appendix

A. GitHub Link

https://github.com/ruchithakor/MRP_Optimal_Biopsy_Decision_Making_Breast_Cancer_RL