Integration of Multivariate Beta-based Hidden Markov Models and Support Vector Machines with Medical Applications

Narges Manouchehri, Nizar Bouguila
Concordia Institute for Information Systems Engineering, Concordia University, Canada
Concordia University, Montreal, Quebec, Canada - H3G 1M8
narges.manouchehri@mail.concordia.ca; nizar.bouguila@concordia.ca

Abstract
In this paper, we propose a novel hybrid discriminative generative model by integrating a modified version of hidden Markov model (HMM), multivariate Beta-based HMM with support vector machine (SVM). We apply Fisher Kernel to define decision boundary and separate classes. In this model, we assume that HMM emission probabilities follow a Beta mixture model as generalizing the assumption of Gaussianity may not be practical in modeling real-world applications. HMM as a generative model needs less amount of data however, its accuracy is less than discriminative models such as SVM. Moreover, in some applications, data may have various feature-length. We tackle this problem with Fisher Kernel. We apply our proposed model to medical applications, lung cancer detection, colonoscopy image, and colon tissue analysis. The results indicate that our proposed model could be a promising alternative.

Introduction
Modeling temporal data is one of the interesting domains of research. Hidden Markov Models as a powerful solution have been used to extract information from dynamic observations. HMM is an attractive solution in healthcare. For instance, it was employed to monitor health and well-being, observe progression of a disease or treatment effects over a specific period. In conventional forms of HMM, it is assumed that emission probabilities follow the Gaussian distribution. However, this belief could not be generalized. In our work, we will suppose that emission probabilities are raised from a finite multivariate Beta mixture model. Considerable flexibility of this distribution in modeling skewed, asymmetric, or symmetric shapes inspired us to choose this option (Manouchehri, Bouguila, and Fan 2021b; Manouchehri, Kalra, and Bouguila 2021; Manouchehri, Bouguila, and Fan 2021c; 2021a). We call our novel modified HMM, multivariate Beta-based HMM (MB-HMM). With the help of such generative models, we can find the underlying pattern by discovering the distribution in each cluster. These models are less accurate compared to discriminative alternatives where we define a hyperplane in multidimensional space to differentiate observations. Support vector machine is a well-known and powerful discriminative classifier and has been applied in numerous medical domains such as cancer prediction or analysis (Huang et al. 2019; Wang et al. 2018), health monitoring (Harimoorthy and Thangavelu 2021) and chronic disease classification (Jain and Singh 2021). Nevertheless, there are always some challenges while dealing with this well-established discriminator such as determining proper kernels. There are some conventional choices such as linear or polynomial ones. But, they have some drawbacks. For instance, data points should have the same feature-length. We could overcome such limitations by choosing Fisher kernels. In this paper, we construct a novel hybrid discriminative generative model by integrating multivariate Beta-based HMM into SVM with Fisher kernels. This new method can capture the pattern of data points with various nature and feature lengths. We will test our model on three real-world medical applications: lung cancer detection, colonoscopy image analysis, and colon tissue analysis. Our main purpose to work on health-related applications is that medical annotation is costly and time-taking. So, we can not easily apply algorithms that need extensive labeling for training. Also, publicly available medical datasets are limited and data-hungry algorithms may not be widely applicable (Litjens et al. 2017). In medicine, it is important to decide based on predictable and explainable results. Thus, some platforms known as black boxes may not be adequate alternatives (Vellido 2020; Gunning 2017; Miotto et al. 2018). Our proposed model is capable enough to model datasets of various sizes and results are explainable in human terms. We compared our model with other alternatives and measure their performance with four metrics: accuracy, precision, recall and F1-score.

This paper is organized as follows: In the section of model specification, we explain the construction of our novel model. We describe first hidden Markov models and extend it to a modified version, multivariate Beta-based hidden Markov models. Afterward, we will discuss forward-backward algorithm as a conventional recursive solution to efficiently evaluate model parameters. In the next step, we will discuss Fisher kernel and integrate discriminative and generative approaches. Then, we present experimental results. Finally, we conclude and discuss our future work.
Model specification

Hidden Markov Models

In first-order hidden Markov model, we suppose that each particular event $t$ is dependent to the immediately previous state at $t - 1$. We formulate a HMM by:

- $\Theta = \{A, B, \alpha, \pi\}$: the complete set of parameters.
- $O = \{O_1, \ldots, O_T\}$: a set of sequential observations. $T$ is the length of sequence.
- $S = \{s_1, \ldots, s_T\}$: the sequence of hidden states such that $s_i \in [1, K]$ and $K$ is the possible number of states.
- Transition matrix: $A = \{a_{ii'} = P(s_t = i' | s_{t-1} = i)\}$ where $\sum_i a_{ii'} = 1$.
- Emission matrix $B = \{b_{ji} = P(o_t = j | s_i = i)\}$ is emission matrix indicating the probability of observing $o_t$ from state $i$ for $j \in [1, M]$.
- $\alpha$ is the shape parameter of MB distribution.
- $\pi = (\pi_1, \ldots, \pi_K)$: Initial values indicating the probabilities to start the sequence from the state $i$. The sum of initial probabilities is 1.

In HMM, we are interested to answer three following questions and find a method to tackle these key issues:

1. What is $P(O | \Theta)$, which is the probability that model $\Theta$ may generate $O$ as a sequence of observation?
2. Assuming that we observed $O$, which sequence of states $S$ could best explain $O$?
3. Given sequential observed data $O$, which model $\Theta$ could be the best to describe them?

To find $P(O | \Theta)$, we calculate the probability of $O$ regardless of a specific sequence of states. So, we take to account all possible states and marginalize over all of them:

$$P(O | \Theta) = \sum_{S} P(S | \Theta) P(O | S, \Theta) \quad (1)$$

$P(O | S, \Theta)$ is the probability of observing a sequence of observation given a set of model parameters:

$$P(S | \Theta) = \pi_{s_1} a_{s_1 s_2} a_{s_2 s_3} \ldots a_{s_{T-1} s_T} = \quad (2)$$

$$p(s_1 | \pi) \prod_{t=2}^{T} p(s_t | s_{t-1}, A)$$

$P(O | S, \Theta)$ defines the probability of observing sequential data $O$ given the sequence of states $S$ is:

$$P(O | S, \Theta) = \prod_{t=1}^{T} P(O_t | s_t, \Theta) = \quad (3)$$

$$b_{s_1 m_1} (O_1) p(O_1 | \alpha_{s_1 m_1}) b_{s_2 m_2} (O_2) p(O_2 | \alpha_{s_2 m_2}) \ldots b_{s_T m_T} (O_T) p(O_T | \alpha_{s_T m_T})$$

We can rewrite $P(O | \Theta)$ as follows:

$$p(O, S, M | A, B, \pi, \alpha) = \quad (4)$$

$$p(s_1 | \pi) \prod_{t=2}^{T} p(s_t | s_{t-1}, A) \prod_{t=1}^{T} p(m_t | s_t, B) p(o_t | \alpha_{s_t m_t})$$

$$\sum_{S} \sum_{L} \pi_{s_1} \prod_{t=2}^{T} a_{s_t s_{t+1}} \prod_{t=1}^{T} b_{s_t m_t} p(o_t | \alpha_{s_t m_t})$$

As we mentioned before, in our work, $p(o_t | \alpha_{s_t m_t})$ is multivariate Beta distribution. To model more than one observation, we sum over the whole sequence. In the next step, we discuss MB distribution in more detail.

Multivariate Beta distribution

We consider $\vec{O} = (o_1, \ldots, o_D)$ as a $D$-dimensional observation. All its features are between zero and one. MB distribution (Olkin and Liu 2003) is described by following probability density function:

$$MB(\vec{O} | \vec{\alpha}) = \frac{\Gamma(\lVert \vec{\alpha} \rVert) \prod_{d=1}^{D} o_{d}^{\alpha_{d}-1} \left[1 + \sum_{d=1}^{D} \frac{o_{d}}{1-o_{d}} \right]^{-|\vec{\alpha}|}}{\prod_{d=0}^{D} \Gamma(\alpha_{d}) \prod_{d=1}^{D} (1-o_{d})^{(\alpha_{d}+1)}} \quad (5)$$

$\vec{\alpha} = (\alpha_0, \ldots, \alpha_D)$ is shape parameter such that $\alpha_d > 0$ for $d = 0, \ldots, D$, $|\vec{\alpha}| = \sum_{d=0}^{D} \alpha_d$ and $\Gamma(\cdot)$ represents the Gamma function. Figure 1 illustrates some samples of this distributions and its mixture models (Manouchehri, Bouguila, and Fan 2021a; 2021c; Manouchehri and Bouguila 2021; Amirkhani, Manouchehri, and Bouguila 2021).

Forward-Backward Algorithm

Considering that we may have several states, various lengths of observations, and numerous data points, calculating (4) may not be feasible directly due to high computational complexity. Inspired by (Rabiner 1989), we apply a method called Forward-Backward algorithm. To use this method, we introduce two assisting functions $f_t(i)$ and $q_t(i)$ which helps us in reducing calculations. In the forward step, we have:

![Figure 1: Multivariate Beta distribution and its mixture models.](image-url)
\[ f_t(i) = p(O_1, \ldots, O_t, s_t = s_i \mid \Theta) \quad (6) \]

It is worthy to mention that \( t \) and \( T \) are different. By calculating \( f(i) \) up to time \( t \), we are reducing the time and it ends up to one specific state at time \( t \) not all possible ones. To solve this equation inductively, we start with a base case for \( O_1 \):

\[ f_1(i) = \pi_i MB_i(O_1), \quad 1 \leq i \leq K \quad (7) \]

Then, by extending this case in inductive step, we have:

\[ f_{t+1}(i') = \sum_{i=1}^{K} f_t(i) a_{ii'} MB_{i'}(O_{t+1}) \quad (8) \]

where \( 1 \leq i' \leq K, 1 \leq t \leq T - 1 \) and at the end we have:

\[ P(O \mid \Theta) = \sum_{i=1}^{K} f_T(i) \quad (9) \]

In backward step, we introduce \( g_t(i) \):

\[ g_t(i) = P(O_{t+1}, O_{t+2}, \ldots, O_T \mid s_t = s_i, \Theta) \quad (10) \]

Similar to forward algorithm, we have a base case:

\[ g_T(i) = 1, \quad 1 \leq i \leq K \quad (11) \]

As the extension, the inductive step is defined by:

\[ g_t(i) = \sum_{i'=1}^{K} a_{ii'} MB_{i'}(O_{t+1}) g_{t+1}(i') \quad (12) \]

for \( 1 \leq i \leq K, t = T - 1, T - 2, \ldots, 1 \).

By combining forward and backward algorithm, we have:

\[ P(O \mid \Theta) = \sum_{i=1}^{K} \sum_{i'=1}^{K} f_t(i) a_{ii'} MB_{i'}(O_{t+1}) g_{t+1}(i') \quad (13) \]

**Fisher Kernels**

Fisher kernel (FK) is a well-known function that has been applied to measure the similarity between two observations. FK lets us combine the positive characteristics of generative and discriminative statistical models. FK is described as follows where FS is Fisher Score:

\[ FK(O, O') = \langle FS(O, \Theta), FS(O', \Theta) \rangle \quad (14) \]

\[ FS(O, \Theta) \] is the gradient of the log-likelihood function with respect to \( \Theta \). So,

\[ FS(O, \Theta) = \nabla_{\Theta} \ln P(O \mid \Theta) \quad (15) \]

The log likelihood of \( P(O \mid \Theta) \) is expressed by:

\[ L(O \mid \Theta) = \ln P(O \mid \Theta) = \ln \sum_{i=1}^{K} \rho_T(i) \quad (16) \]

\[ = \ln \sum_{i=1}^{K} \pi_i MB_i(O_1) g_t(i) \]

We compute the derivatives of log-likelihood with respect to \( \Theta \) as follows:

\[ \nabla_{\Theta} L(O \mid \Theta) = \left[ \frac{\partial L(O \mid \Theta)}{\partial \pi_i}, \frac{\partial L(O \mid \Theta)}{\partial a_{ii'}}, \frac{\partial L(O \mid \Theta)}{\partial \alpha_{id}} \right] \quad (17) \]

\[ \frac{\partial L(O \mid \Theta)}{\partial \pi_i} = MB_i(O_1) g_t(i) \sum_{i=1}^{K} \pi_i MB_i(O_1) g_t(i) \quad (18) \]

\[ \frac{\partial L(O \mid \Theta)}{\partial a_{ii'}} = \frac{1}{P(O \mid \Theta)} \sum_{k=1}^{K} \frac{\partial \rho_T(k)}{\partial a_{ii'}} = (19) \]

\[ \frac{1}{P(O \mid \Theta)} \sum_{k=1}^{K} \left( \frac{\partial}{\partial a_{ii'}} \sum_{l=1}^{K} \rho_T-l(b_{lk} MB_k(O_T)) \right) = \quad (20) \]

\[ \frac{\partial L(O \mid \Theta)}{\partial \alpha_{id}} = \frac{1}{P(O \mid \Theta)} \times \left( \sum_{k=1}^{K} \sum_{l=1}^{K} \frac{\partial \rho_T-l(k)}{\partial \alpha_{id}} b_{lk} MB'_k(O_T) + \sum_{k=1}^{K} \rho_T-l(k) b_{ki} \frac{\partial MB_i(O_T)}{\partial \alpha_{id}} \right) \quad (21) \]

\[ \frac{\partial MB_i(O_1)}{\partial \alpha_{id}} = \Psi \left( \sum_{d=0}^{D} \alpha_{id} \right) \quad (22) \]

\[ - \Psi(\alpha_{id}) + \ln o_d - \ln(1 - o_d) - \ln \left[ 1 + \sum_{i=1}^{k} \frac{o_d}{1 - o_d} \right] \quad (23) \]

\( \Psi(\_) \) is digamma function as the first derivative of the Gamma function.

**Experimental Results**

To validate the performance of our proposed hybrid method, we test it on three real medical applications: lung cancer detection, colon histopathology, and endoscopic image analysis. We compare four models including hybrid multivariate Beta-based HMM with SVM (MB-HMM-SMV), finite multivariate Beta-based HMM (MB-HMM), Gaussian...
mixture-based HMM (GMM-HMM), and SVM with the linear kernel (SVM-linear). We extracted features of images with VGG-16 and considering the nature of our model, we need to normalized data with min-max method. We consider four following metrics to assess models robustness where $TP, TN, FP$ and $FN$ are the total number of true positives, true negatives, false positives, and false negatives, respectively:

$$\text{Accuracy} = \frac{TP + TN}{\text{Total number of observations}} \quad (22)$$

$$\text{Precision} = \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN}$$

$$F1 - \text{score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

### Lung tissue analysis

Lung cancer is one of the leading roots of global mortality. Based on the statistics released by World Health Organization (WHO) in 2020, lung cancer is the second most common type of cancer with 2.21 million diagnosed cases in the world and has the first place among the most common reasons for cancer-related death with 1.80 million deaths (WHO 2020). This disease could be classified into three categories: carcinoid, small, and non-small cell lung cancers. Analyzing histopathology images to differentiate the cases based on their morphological characteristics and cellular appearance is a principle method. This procedure is strongly important as it directly affects the decision of physicians and has a high impact on patients’ life. Any error in diagnosis may result in irreversible side effects. So, timely and accurate analysis of histopathology images is vital. Computer-assisted algorithms and machine learning algorithms could help physicians as a second-opinion method. In this work, we used a publicly available dataset (Borkowski et al. 2019) with 750 cases consisting of 250 cases in three groups: benign, adenocarcinoma, and squamous cell carcinomas lung tissue. We augmented the size of this dataset to have 1000 samples in each cluster. Some samples are shown in Figure 2. We present the results of model evaluation in Figure 3. As it is shown, our proposed hybrid method MB-HMM-SVM provides the best results with 87.44%, 87.51%, 88.07%, and 87.79% of accuracy, precision, recall, and F1-score, respectively. Conventional discriminative methods such as SVM with linear kernel or generative algorithms such as GMM-HMM which are widely used have less accurate results. These outcomes encourage us to consider our novel method as a capable alternative.

### Colonoscopic image analysis

Considering the reports of WHO, colon and rectum cancer has the third rank among the most common cancers in 2020. With 1.93 million diagnosed cases and 935,000 mortality reports (WHO 2018), this disease has grabbed a considerable amount of attention. There are some main causes such as colorectal polyps or family history which play important roles in monitoring, prevention, and early diagnosis of this disease. In the majority of cases, benign polyps and precursor lesions are formed which are the starting step of cancer. If remained untreated, the texture of benign polyps will change over time. So, screening and starting to take preventive actions are important. Such policies are more vital for patients with inherited gene mutations. However, with early detection survival rate is increased. Colonoscopy is the gold standard technique and a primary screening test for detecting colon abnormalities and following up the progress of diseases after treatment. This minimally invasive procedure, help the gastroenterologists to find and remove polyps (Zauber et al. 2012; Winawer et al. 1993). However, the complicated anatomy of colon, and various shapes of polyps are some challenging issues in this procedure (Tholoor et al. 2013). Similar to other fields of science, machine learning could help us to reduce false-diagnosed or missed cases (Kumar et al. 2017; Rees et al. 2017). We tested our novel method on a publicly available dataset (Zauber et al. 2012) which includes 500 normal colon and 500 polypoid lesions images. In Figure 4, we illustrate some of these endoscopic images. The outcomes of evaluation are shown in Figure 5. MB-HMM-SVM has the best results considering four metrics, accuracy, precision, recall, and F1-score with 93.81%, 93.35%, 92.02%, and 92.68%, respectively. MB-HMM is the second-ranked with 87.26% of F1-score. Similar to the previous experiment, GMM-HMM and SVM with linear kernel have less promising performance with 81.37 and 77.07 percent of F1-score.

![Figure 2: Examples of adenocarcinoma, benign, and squamous cell carcinomas lung tissues from left to right of image.](image)

![Figure 3: Results of model evaluation on lung dataset.](image)
Colon histopathology tissue analysis

As we discussed in the previous experiment, colorectal carcinoma is one of the frequent types of cancer. However, unlike many others, colorectal cancer (CRC) is preventable. During colonoscopy, as the most precise test, physicians examine the colon, resect polyps and take some biopsies from tissues or polyps that they have doubts about. The next step is analyzing the specimens by pathologists to find out the histopathology and molecular structure of colon tissue. This phase is crucial as it will affect disease management and prognosis assessment. Adenocarcinoma is conventionally distinguished by glandular formation as the beginning phase in tumor grading. In Figure 6, some examples of benign and malignant cases are illustrated. Differentiating the tissues by human eyes could be prone to errors. Machine learning-based computational pathology could be an emerging and promising discipline that can augment the accuracy of clinical decisions. This results in increasing the quality of services provided by healthcare systems. In this part of our experiment, we applied our proposed model on a publicly available dataset (Borkowski et al. 2019). We have two groups of samples, benign and malignant tissues with 1000 cases in each class. Referring to the results shown in Figure 7, MB-HMM-SVM with 91.28% of F1-score provides the best results. This output is followed by the outputs of MB-HMM with 89.56% of F1-score. Similar to two previous experiments, GMM-HMM and conventional SVM with linear kernel did not have good results compared to the two other models.

Conclusion

In this paper, we have proposed a new hybrid discriminative-generative method by combining a novel HMM with SVM. We validated this newly proposed model using three challenging medical applications. Integrating two powerful machine learning methods assist us in having a promising model. Also, we will be able to model observations of various lengths. This makes our model a proper choice for real-world applications. In multivariate Beta-based HMM as the generative part of our model, we assume that emission probabilities are following MB mixture. This relaxes the constraint of Gaussianity and provides more flexibility to fit numerous shapes of data. Fisher Kernel in SVM aids in modeling data in realistic scenarios. We applied our model to detect cancer of lung and colon by histopathological samples. Also, we studied how our proposed method could help physicians in finding colon polyps. We compared the performance of our novel algorithm with conventional alternatives such as GMM-HMM and SVM with linear kernel. Considering the results, we realize that our proposed method is a robust approach and could be considered as a promising technique compared to widely-used methods. Our future research could be devoted to integrating other flexible distributions into hidden Markov models and combining them with discriminative tools.
Acknowledgment
The completion of this research was made possible thanks to the Natural Sciences and Engineering Research Council of Canada (NSERC).

References
Amirkhani, M.; Manouchehri, N.; and Bouguila, N. 2021. Birth-death MCMC approach for multivariate beta mixture models in medical applications. In International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, 285–296. Springer.

Borkowski, A. A.; Bui, M. M.; Thomas, L. B.; Wilson, C. P.; DeLand, L. A.; and Masterides, S. M. 2019. Lung and colon cancer histopathological image dataset (lc25000). arXiv preprint arXiv:1912.12142.

Gunning, D. 2017. Explainable artificial intelligence (xai). Defense Advanced Research Projects Agency (DARPA), nd Web 2.

Harimooorthy, K., and Thangavelu, M. 2021. Multi-disease prediction model using improved svm-radial bias technique in healthcare monitoring system. Journal of Ambient Intelligence and Humanized Computing 12(3):3715–3723.

Huang, H.; Zhou, S.; Jiang, J.; Chen, H.; Li, Y.; Li, C.; et al. 2019. A new fruit fly optimization algorithm enhanced support vector machine for diagnosis of breast cancer based on high-level features. BMC bioinformatics 20(8):1–14.

Jain, D., and Singh, V. 2021. A two-phase hybrid approach using feature selection and adaptive svm for chronic disease classification. International Journal of Computers and Applications 43(6):524–536.

Kumar, S.; Thosani, N.; Ladabaum, U.; Friedland, S.; Chen, A. M.; Kochar, R.; and Banerjee, S. 2017. Adenoma miss rates associated with a 3-minute versus 6-minute colonoscopy withdrawal time: a prospective, randomized trial. Gastrointestinal endoscopy 85:1273–1280.

Litjens, G.; Kooi, T.; Bejnordi, B. E.; Setio, A. A. A.; Ciompi, F.; Ghafoorian, M.; Van Der Laak, J. A.; Van Ginneken, B.; and Sánchez, C. I. 2017. A survey on deep learning in medical image analysis. Medical image analysis 42:60–88.

Manouchehri, N., and Bouguila, N. 2021. Stochastic expectation propagation learning of infinite multivariate beta mixture models for human tissue analysis. In IECON 2021–47th Annual Conference of the IEEE Industrial Electronics Society, 1–6. IEEE.

Manouchehri, N.; Bouguila, N.; and Fan, W. 2021a. Batch and online variational learning of hierarchical dirichlet process mixtures of multivariate beta distributions in medical applications. Pattern Analysis and Applications 24(4):1731–1744.

Manouchehri, N.; Bouguila, N.; and Fan, W. 2021b. Batch and online variational learning of hierarchical pitman-yor mixtures of multivariate beta distributions. In 20th IEEE International Conference on Machine Learning and Applications, ICMLA 2021. Pasadena, CA, USA, December 13-16, 2021, 298–303. IEEE.

Manouchehri, N.; Bouguila, N.; and Fan, W. 2021c. Non-parametric variational learning of multivariate beta mixture models in medical applications. International Journal of Imaging Systems and Technology 31(1):128–140.

Manouchehri, N.; Kalra, M.; and Bouguila, N. 2021. Online variational inference on finite multivariate beta mixture models for medical applications. IET Image Processing.

Miotto, R.; Wang, F.; Wang, S.; Jiang, X.; and Dudley, J. T. 2018. Deep learning for healthcare: review, opportunities and challenges. Briefings in bioinformatics 19(6):1236–1246.

Olin, I., and Liu, R. 2003. A bivariate beta distribution. Statistics & Probability Letters 62(4):407–412.

Rabiner, L. R. 1989. A tutorial on hidden markov models and selected applications in speech recognition. Proceedings of the IEEE 77(2):257–286.

Rees, C. J.; Rajasekhar, P. T.; Wilson, A.; Close, H.; Rutter, M. D.; Saunders, B. P.; East, J. E.;Maier, R.; Moorghen, M.; Muhammad, U.; et al. 2017. Narrow band imaging optical diagnosis of small colorectal polyps in routine clinical practice: the detect inspect characterise resect and discard 2 (discard 2) study. Gut 66(5):887–895.

Tholoor, S.; Tsagkournis, O.; Basford, P.; and Bhandari, P. 2013. Managing difficult polyps: techniques and pitfalls. Annals of Gastroenterology: Quarterly Publication of the Hellenic Society of Gastroenterology 26(2):114.

Vellido, A. 2020. The importance of interpretability and visualization in machine learning for applications in medicine and health care. Neural Comput. Appl. 32(24):18069–18083.

Wang, H.; Zheng, B.; Yoon, S. W.; and Ko, H. S. 2018. A support vector machine-based ensemble algorithm for breast cancer diagnosis. European Journal of Operational Research 267(2):687–699.

WHO. 2018. Colorectal cancer report of World Health Organization (WHO). https://www.who.int/news-room/fact-sheets/detail/cancer.

WHO. 2020. Lung cancer report of World Health Organization (WHO). https://www.who.int/news-room/fact-sheets/detail/cancer.

Winawer, S. J.; Zauber, A. G.; Ho, M. N.; O’Brien, M. J.; Gottlieb, L. S.; Sternberg, S. S.; Waye, J. D.; Schapiro, M.; Bond, J. H.; Panish, J. F.; et al. 1993. Prevention of colorectal cancer by colonoscopic polypectomy. New England Journal of Medicine 329(27):1977–1981.

Winawer, S. J.; Zauber, A. G.; Ho, M. N.; O’Brien, M. J.; Lansorp-Vogelaar, I.; van Ballegooijen, M.; Hankey, B. F.; Shi, W.; Bond, J. H.; Schapiro, M.; Panish, J. F.; et al. 2012. Colono- scopic polypectomy and long-term prevention of colorectal-cancer deaths. N Engl J Med 366:687–696.