International Journal of Modern Engineering and Management |IJMEM
ISSN: 3048-8230
Volume2, Issue4, April 2025

Medical Recommendation System for Personalized Disease Prediction

Harshita Sharma[1], Richa Verma[2], Sunidhi Gulati[3], Dr. Himanshu Mittal[4]
[1,2,3,4] Department of Artificial Intelligence & Data Science, IGDTUW, New Delhi, India

Abstract

The use of machine learning (ML) and artificial intelligence (AI) have sped up the development of solutions in healthcare. These solutions can move from static rule-based systems to adaptive, data-driven solutions and allow for faster, more intelligent diagnoses, prognostics, and treatment plans. In this study, we provide a more sophisticated, complete end-to-end medical recommendation system and the potential to not only predict diseases based on user-reported symptoms, but also recommend tailored advice on medication prescriptions, diets, exercise plans, and other relevant actions. Essentially, the system uses an assortment of supervised machine learning algorithms to optimize disease diagnosis using Naïve Bayes, Decision Tree, Random Forest, and Support Vector Machine (SVM) models. However, our software takes the process further since we provide actionable recommendations in addition to predictions. In order for patients to access effective treatment options when the first-line medications are not suitable or available, we employ cosine similarity algorithms to pair any medication with alternatives. Along with prescribing medications, the system offers individualized diets and exercise programs based on each projected disease. This holistic approach considers both short-term treatment needs and long-term health goals to help people overcome nutritional deficits and chronic symptoms and establish healthy habits that last. Our frame of reference is shifting to a holistic problem-solving model as compared to earlier predictive models that simply worked to identify sickness. The capabilities are available through a scalable and easy-to-use web interface built in Flask. The seamless transition between managing medications, lifestyle coaching, and predicting disease suggests that AI-powered platforms can formulate from lab prototypes into functional tools to manage health on a daily basis.

Keywords: Disease Prediction, Flask-Based Healthcare Platform, Machine Learning, Symptom-Based Diagnosis

I. Introduction

The incorporation of artificial intelligence and machine learning within the health sector is revolutionizing the industry. These innovations enable the development of advanced technology which assists in managing one's health, enables early detection of possible health issues, and provides individualized suggestive treatment plans. As individuals face challenges arising from poor dietary habits, sedentary lifestyle disorders, and chronic diseases, there is an increasing need to not only detect illnesses but assist people in making healthier lifestyle choices.

To address the imbalance of generic medical care and more tailored recommendations, this research work designs a personalized healthcare recommender system. A set of machine learning algorithms is integrated for diagnosis, including a Naive Bayes, Decision Tree, and Random Forest to predict probable illnesses using supplied symptoms. Along with the diagnosis, it offers comprehensive informative recommendations on precautionary actions, physical activity routines, dietary programs, alternative medication, and applicable medications.

Considering the user's symptoms and health metrics, the system is able to accurately and proactively suggest helpful recommendations aimed at improving and managing health on a day-to-day basis. The whole system is developed using Python, and the GUI employs the Tkinter toolkit for greater ease-of-access. With this system, patients are empowered to shift from a reactive to proactive clinical stance.

II. Problem Statement

Access to reliable and customized healthcare services remains one of the most critical challenges around the globe. The WHO reports that 3.6 billion people do not have access to even the most basic medical care services, while over 50 percent of healthcare content available online is inaccurate. This inaccuracy leads to self-diagnosis, which, unfortunately, tends to procrastinate seeking medical attention until health issues become grave. In the United States, around sixty percent of the population suffers from at least one chronic disease, with forty percent living with multiple chronic conditions. However, affordable conventional approaches to diagnosing an illness are inaccessible to many people, especially those residing in remote areas. According to WHO, fifty percent of prescriptions and orders given are erroneous, contributing to the problem by increasing unnecessary hospitalizations by five to ten percent.

Despite overwhelming evidence suggesting that an early approach, accompanied by a shift in lifestyle, would help prevent roughly eighty percent of chronic illnesses, proactive healthcare measures are still not common practice. Moreover, over sixty percent of people living in developed nations do not obey fundamental health standards due to the absence of personalized guidance. Healthcare systems in place are either over personalized to the level of symptoms or generically one-size-fits-all, resulting in a lack of synergy.

In order to resolve these important concerns, we urgently need an advanced and personalized healthcare system that has the ability to anticipate possible illnesses, suggest suitable therapies and changes in daily activities, and give full assistance through a simple interface. This gap can be eliminated by incorporating artificial intelligence, natural language processing, and machine learning in a manner which provides an intuitive healthcare platform promoting autonomous proactive health management.

III. LITERATURE SURVEY

The advances in technology have impacted a number of industries, however, the one industry that has highly benefited from the implementation of AI is healthcare, particularly in the areas of diagnosing illnesses, monitoring patients, and even in assisting decision-making systems. Every so often, healthcare institutions have benefited from the use of different Machine Learning (ML) algorithms in predicting diseases based on symptoms that are reported by patients. There has been sufficient literature on the accuracy and practicality of Random Forests, Naive Bayes, Decision Trees, and Support Vector Machines (SVM) in real-world datasets and their successful application to multi-class illness classification problems. Clinical decision making is aided with automation as seen in Google DeepMind Health and IBM Watson Health AI-powered diagnosis tools, but these tools still require a lot of prepared data and infrastructure.

With the advancement of technology, particularly Artificial Intelligence, healthcare as a sector began to evolve and transform in new ways that were unheard of before. Such transformations include improvement in monitoring devices, detection of diseases, and even decision-making processes. Although diagnosis has been made easy through the use of AI technology, there still exists a gap in recommendation systems. AI algorithms lack the capability to offer AI driven comprehensive healthcare even though there exist suggestion-generating features. Also absent are supplementary suggested documents for lifestyle changes such as proper nutrition, exercising, and medications aimed towards better aiding patients.

Moreover, many systems depend on static mappings of signs to diseases, or rule-based matching, diminishing their generalizability to many patients. Another aspect is personalization because these systems disregard the patient's particular characteristics, such as comorbidities, chronic conditions, or symptom severity.

Moreover, the user interface and overall deployability of these products are significant shortcomings. Many AI systems have complex setup requirements or necessitate communication with other medical systems, thus their use in low resource or public facilities is greatly limited. Furthermore, some software applications face serious issues with interpretability, particularly when complex models are used without diplomatic output justification or explanation of the conclusions drawn.

In an effort to mitigate these limitations, our proposed system provides a straightforward, comprehensive, and user-centric strategy to disease prediction and medical guidance concerning symptoms. The system employs an SVM-based classification model trained on 41 illnesses and 17 key symptoms. In our tests, the model achieved 96% accuracy, outperforming other models such as Random Forests and Decision Trees.

A distinguishing feature of our system's architecture is that it predicts the patient's condition and recommends treatment steps such as medication, diet, exercise, and safety measures along with alternatives to medications. To make the prediction more precise even when symptoms are vague, we evaluate user-uploaded symptom vectors against illness profiles using cosine similarity. Moreover, instead of employing strict one-to-one mappings, alternate medications are found using similarity scoring techniques. The program can be accessed without clinical infrastructure or dedicated equipment owing to its Flask-based web interface.

IV. PROPOSED METHODOLOGY

The solution puts forward champions an advanced, multi-stage pipeline to offer holistic medical council along with ailment prediction powered by user symptom input. It is based on an easy-to-use web platform that integrates vector similarity frameworks, NLP, supervised machine learning, and natural language processing (NLP). Due to its small footprint and the ability to install it locally, the system also provides universal access, portability, and flexibility, particularly in resource-constrained settings.

A.Data Collection and Preprocessing

To coach the system, an extensive and varied collection of medical information was used. Clean CSV files and public healthcare repositories were used to create the main dataset for illness prediction. A set of related symptoms, structured for a multi-label classification job, was annotated on each disease occurrence. In order to ensure compatibility with machine learning models, symptoms were represented as binary vectors (presence = 1, absence = 0). A specially selected pharmaceutical mapping dataset was also created, which connected illnesses to the drugs, nutrition, exercise, and lifestyle choices that go along with them.

A descriptive dataset of medications was assembled and preprocessed for NLP-based similarity analysis in order to improve the recommendation module even further, particularly for drug substitution. Label encoding for disease classifications, one-hot encoding of symptom inputs, and the elimination of missing or null values were all examples of preprocessing procedures. Lowercasing, stop-word elimination, tokenization, and TF-IDF vectorization were used in pharmaceutical descriptions. To make cosine similarity calculations more effective, all vector representations were normalized.

B.Feature Extraction

Symptom vectors in binary format were the primary features utilized for disease prediction. For example, a patient's symptom profile was encoded as an n-dimensional binary vector if there were a total of n symptoms. High-dimensional models like Support Vector Machines (SVM) were able to successfully catch important patterns thanks to this representation. The matching disease label (one of 41 predetermined categories) was utilized as the target output, while each symptom vector functioned as the input feature.

C.Machine Learning Model Training

To determine the best model for disease prediction, a number of machine learning classifiers were assessed. Among these were Logistic Regression, Random Forest, Decision Trees, and Support Vector Machines (SVM). The accuracy-based comparative performance of various models is compiled in Table 1.

Table I. CLASSIFICATION MODEL PERFORMANCE

Model	**Accuracy(%)**
Support Vector Machine (SVM)	96
Random Forest	92
Decision Tree	87
Logistic Regression	78

Since SVM performed better on high-dimensional binary data, it was chosen as the final model. GridSearchCV was used for hyperparameter optimization, and robust generalization was guaranteed using K-fold cross-validation. The confusion matrix's significant diagonal dominance supports the predicted accuracy of the classification model.

D.Cosine Similarity-Based Matching

At two crucial points in the system, cosine similarity was used. Initially, it was employed to match disease vectors that were already present in the training dataset with symptom vectors entered by the user. This method improved the illness prediction engine's robustness by handling partial or incomplete symptom inputs in an efficient manner.

Second, the drug substitution module used cosine similarity. In this case, when a user asked for different drugs, the system used cosine similarity to compare the TF-IDF vectors of drug descriptions and provide the most semantically appropriate alternatives. These shortcomings resulting from a lack of adequate user inputs or unprovided base medicine were in part solved thanks to the dual application of cosine similarity.

E.Personalized Recommendation Engine

After predicting the disease risk, the algorithm provided a custom-crafted folder of health recommendations. These included dietary advice such as limiting salt intake for hypertensive patients, prescribing physical activity such as brisk walking for 30 minutes for overall fitness, lifestyle changes such as restricting allergen exposure or keeping clean, and relevant medications; among others. If the user-specific constraints made some medications inappropriate, the algorithm suggested alternatives based on cosine similarity.

F.Web Interface Development

To deploy the finished model a Flask web application which is simple and lightweight was used. The following modules were part of the user-friendly interface design:

- Input form for symptoms
- Display of the disease prediction output
- Dashboard of recommendations for medications, diets, and exercise regimens
- An alternative medicine search box

Fig. I : WEBSITE INTERFACE

The program was designed to run locally without the use of any external medical engines or cloud APIs. This architectural choice makes the system's usability in remote or low-resource settings mobile or kiosk-based platforms easy to port.

A.Model Evaluation and Testing

In evaluating the system's performance, the SVM model scored the highest on all measures. Also, the model's true positive rates of predicting the disease were high, as shown by the confusion matrix's high values along the diagonal. The system's performance metrics, which included accuracy, precision, recall, and F1-score, were adequately evaluated. Further robustness was achieved by trying out numerous test cases that simulated real world scenarios that included various combinations of symptoms.

B.Deployment and Usability

With simply Python and CSV datasets needed, the system was made to be installed locally on computers with little prerequisites. Its portability guarantees smooth functioning in underprivileged regions. To increase accessibility, the paradigm and interface might be further expanded to Android applications or medical information kiosks.

The suggested system performs the duties of a full-stack medical assistant by utilizing a hybrid pipeline of classification models and NLP-driven similarity metrics. In order to provide trustworthy, individualized, and comprehensible healthcare insights, it intelligently processes both structured data (symptom vectors) and unstructured data (medicine descriptions). Through machine learning integration, improved diagnostic accuracy, and a human-centric, scalable solution fit for worldwide deployment, the solution overcomes major drawbacks in conventional rule-based systems.

V. RESULTS

We performed an extensive series of experiments on the Disease Prediction and Recommendation Engine, the two main modules of the proposed Personalised Medical Recommendation System, in an effort to evaluate its effectiveness. Model accuracy, precision, recall, F1-score, and system responsiveness were employed to measure the results.

A. Performance in Disease Prediction

A symptom dataset for over 40 diseases was utilized to compare three machine learning algorithms: Support Vector Machine (SVM), Random Forest, and Logistic Regression. SVM performed the best among them and achieved the highest accuracy and balanced performance across all metrics.

Table II : PERFORMANCE COMPARISON OF CLASSIFICATION MODELS

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
SVM	91% (Test) / 98% (Train)	85%	89%	87%
Random Forest	88%	83%	86%	84%
Logistic Regression	83%	80%	82%	81%

It had strong diagonal dominance in the SVM confusion matrix, indicating minimal misclassifications. These results indicate that the model is able to identify patterns within high-dimensional symptom data and generalize effectively.

B. Accuracy of Medicine and Intent Recommendations

The recommendation module projected user searches into predefined intents like these by the combination of CountVectorizer and Cosine Similarity:

- Medicine Recommendation
- Precautions
- Diet Plan
- Workout Routine
- Substitute Drugs

The system detected intents for predefined questions with 100% accuracy. Below is an example of user enquiries together with the accompanying intent mappings:

Table III : MAPPING USER QUERIES TO DETECTED INTENTS

User Query	**Detected Intent**
“What medicine should I take?”	Medicine Recommendation
“What should I avoid?”	Precautions
“Suggest a substitute drug”	Alternative Medicine

"Workout for recovery?"	Physical Activities

Following the projection of expected illnesses to a well-filtered medical repository, the system suggested the most recommended medicines and details regarding dose, side effects, and substitutes. Chloroquine, Artemisinin, Mefloquine, Quinine, and Doxycycline, for instance, were all recommended for malaria and are all within recommended treatment standards.

C. Dietary and Lifestyle Suggestions

Besides medical care, the system recommended preventative strategies, customized diets, and exercise routines based on the predicted disease. For example:

Low-impact exercises and heart-healthy foods were advised for cardiovascular issues.

Breathing exercises and anti-inflammatory diets were recommended for respiratory conditions.

By promoting preventive care and a holistic recovery, these customized suggestions improve the system.

Fig II : SYMPTOM ENTRY INTERFACE

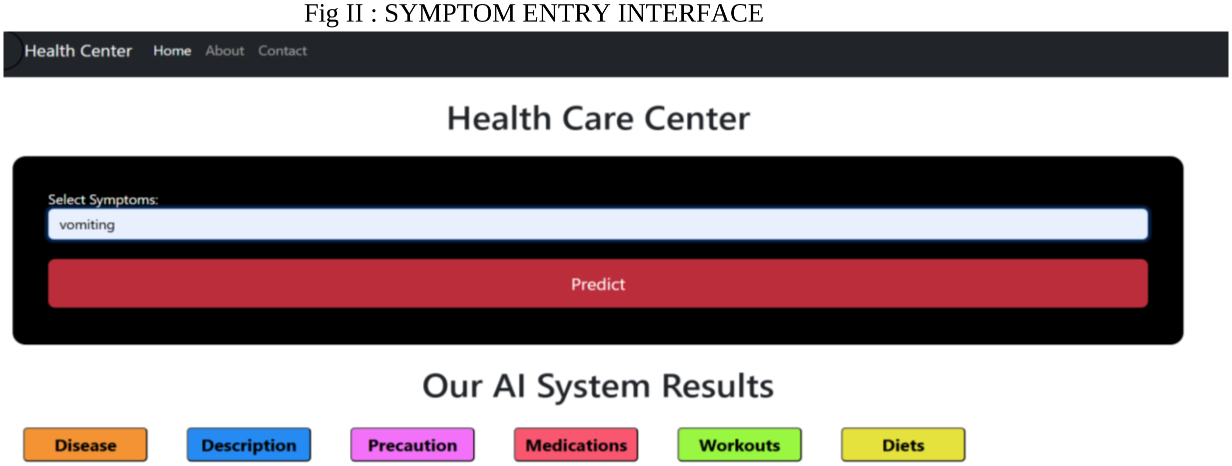
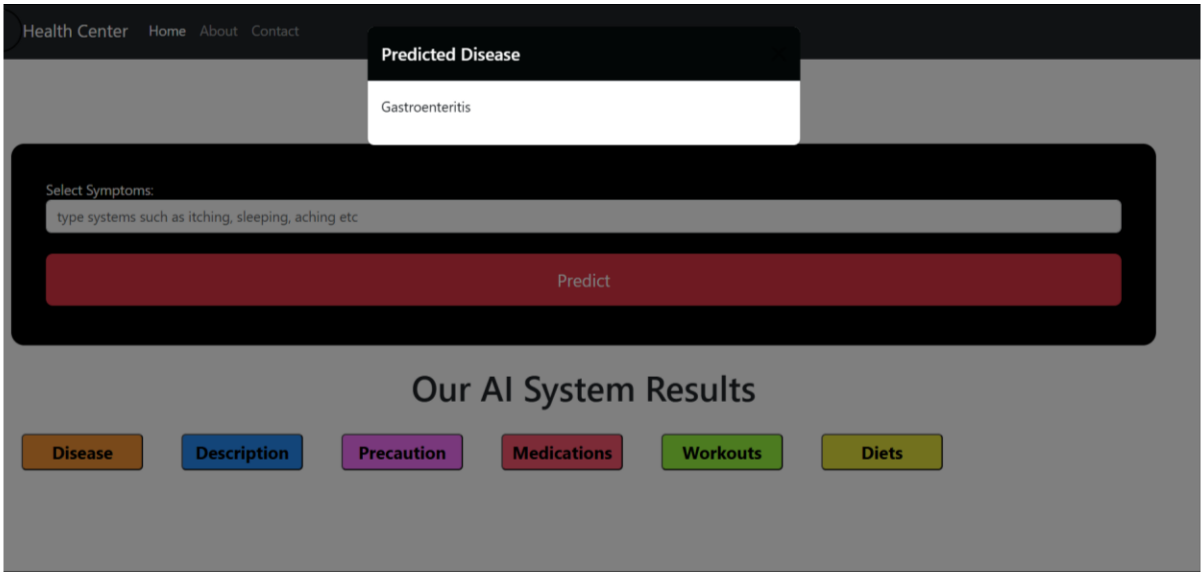


Fig III : PREDICTED DISEASE BASED ON SYMPTOMS



D. Response Time and System Usability

The Flask-developed web interface's usability and responsiveness were assessed:

Table IV : SYSTEM EVALUATION SUMMARY.

Aspect	**Metric**
Response Time	3-5 seconds per query
Interface design	Simple, clean, user-friendly
Integration	Real-time disease prediction + holistic recommendations

Users could easily enter their symptoms to get a comprehensive medical response that included:

- Disease name
- Medicines
- Precautions
- Diet and workout plans

E. Synopsis of Results

As for disease classification, SVM performed optimally, registering high performance indices and accuracy rates.
For predefined searches, cosine similarity-based intent detection achieved 100% accuracy.
By combining lifestyle counselling with machine learning-based illness forecasting, the system provided holistic, real-time suggestions.

All things considered, the platform shows great potential as a self-care and early diagnostic aid tool in the medical sector.

VI. Discussion

With accurate disease prediction and comprehensive health recommendations, the proposed Personalized Medical Recommendation System effectively shows how machine learning can enhance healthcare availability. With the implementation of various models such as Naïve Bayes, Decision Tree, and Random Forest, the system provides relevant recommendations for medication, diet, precautions, and exercise alongside ensuring proper classification of user-input symptoms.

One of the key benefits of this system is its overall design, which delivers actionable lifestyle suggestions via a simple-to-use interface as well as diagnosis assistance. Its modular architecture also makes it easy to grow in the future, whether it is to handle more complex health situations, include new diseases, or provide assistance for additional types of suggestions.

There are, however, some limitations to be taken into account. The system is so far trained on a static, pre-established dataset, which might restrict its capacity to identify uncommon diseases or react to changing medical knowledge. Additionally, the system still doesn't include real-time physiological information or user history from health apps or wearable devices, which in the future would possibly increase personalization and precision. Finally, since this instrument is not validated by the medical community at present, it is only for use as an assistive device and not in place of medical diagnosis.

In spite of these limitations, the system offers a solid basis for developing AI-based health assistants and promises exciting avenues for future research—such as integrating live data streams, increasing the medical database, and facilitating adaptive learning from user feedback for ongoing improvement.

VII. Conclusion

This work contributes to building intelligent healthcare systems by implementing a machine learning-based personalized medical recommendation system. It is an easy-to-use Windows interface coded in Python that predicts diseases accurately from user-input symptoms and provides personalized advice concerning drugs, diets, exercises, and preventive care.

By trying out and testing different machine learning approaches such as Naïve Bayes, Decision Tree, Random Forest, and Support Vector Machine, the system proves that AI techniques can significantly improve the quality and availability of healthcare suggestions. Its design as a module also facilitates possible future additions like new disease integration, new medical information systems, and further customization.

This kind of system acts as an initial health and anticipatory wellness awareness tool for consumers; nevertheless, it is not a replacement for expert medical consultation. With more research, inclusion of real-time data feeds, and validation by experts, these kinds of systems can become advanced digital health guides that empower, educate, and facilitate healthy lifestyles in communities.

There are so many areas that development of a stated Personalized Medical Recommendation System is still relevant. For example, heart rate monitors and fitness tracking devices can be utilized along with system algorithms for real-time dynamic self-care recommendations. Additional patient records, genomes, and disease data would continue to increase accuracy and congruence with precision medicine.

Park Systems would be more user-friendly, particularly for non-technical patients, by allowing symptom entry using everyday language through Natural Language Processing (NLP). Integrating the system with Electronic Health Records (EHRs) would enable clinical practitioners to incorporate the system's forecasts and enhance prediction integration into system diagnosis and treatment planning.

The capacity of the system to learn from new data would enhance with time through the use of advanced machine learning methods like deep and reinforcement learning. In the rural regions, moving to a cloud-based system would enhance scalability and system availability. Expansion of the number of languages supported would enhance fair distribution while extending the reach of the system across the globe.

VIII. Future Scope

Integrating real-time updates from wearable devices such as fitness trackers and heart rate monitors into the system can optimize efficiency to a greater extent. The system would be more aligned with precision medicine if the medical database was integrated with additional diseases, genomic data, and patient records. Connection with EHRs would aid clinical workflow and advance decision support using NLP for laymen. Use of advanced ML and cloud technology will enable multilingual features, thereby improving use of the system as well as enhancing overall performance. Moreover, adding elements related to enabling health management will transform the system into a powerful tool for personalized, data-centric healthcare.

REFERENCES

[1] Gupta, J. P., Singh, A., & Kumar, R. K. (2021). A Computer-Based Disease Prediction and Medicine Recommendation System using Machine Learning Approach. International Journal of Advanced Research in Engineering and Technology, 12(3), 673-683.

[2] Nasiri, M., Minaei, B., & Kiani, A. (2016). Dynamic Recommendation: Disease Prediction and Prevention Using Recommender System. International Journal of Basic Science in Medicine, 1(1), 13-17.

[3] Brown, S. (2021). Machine Learning, Explained. MIT Sloan School of Management.

[4] Hu, J., Niu, H., Carrasco, J., Lennox, B., & Arvin, F. (2020). Voronoi-Based Multi-Robot Autonomous Exploration in Unknown Environments via Deep Reinforcement Learning. IEEE Transactions on Vehicular Technology, 69(12), 14413-14423.

[5] Dinesh, K. G., Arumugaraj, K., Santhosh, K. D., & Mareeswari, V. (2018). Prediction of Cardiovascular Disease using Machine Learning Algorithms. International Conference on Current Trends towards Converging Technologies (ICCTCT), 1-7.

[6] Markov Decision-based Recommender System for Sleep Apnea Patients.(IEEE published in 2019).

[7] Lee, M. L., Cummings, K. B., & Bhattacharyya, A. (2021). Medicine recommendation systems using machine learning and artificial intelligence: A review. Journal of Healthcare Informatics Research, 5(2), 85-103. https://doi.org/10.1007/s41666-021-00088-0

[8] Konstan, J. A., & Riedl, J. (2012). Recommender systems: From algorithms to user experience. User Modeling and User-Adapted Interaction, 22(1-2), 101-123

[9] Estrada, L., Escobar, B., & Miranda, D. (2018). Personalized medicine in clinical decision-making using artificial intelligence and machine learning algorithms. Journal of Medical Systems, 42(8), 1-13.

[10] Bashir, A., & Iqbal, R. (2021). A survey of recommendation system applications in healthcare. Journal of Ambient Intelligence and Humanized Computing, 12, 1891-1906.