IBCA: An Intelligent Platform for Social Insurance Benefit Qualification Status Assessment

Yuliang Shi[1,2,3*], Lin Cheng[1], Cheng Jiang[2], Hui Zhang[1,2], Guifeng Li[2], Xiaoli Tang[3,4], Han Yu[3,4*], Zhiqi Shen[3,4*], Cyril Leung[3,5]

[1]School of Software, Shandong University (SDU), Jinan, China
[2]Dareway Software Co. Ltd, Jinan, China
[3]Joint SDU-NTU Centre for Artificial Intelligence Research (C-FAIR), Shandong University (SDU), Jinan, China
[4]School of Computer Science and Engineering, Nanyang Technological University (NTU), Singapore
[5]Department of Electrical and Computer Engineering, The University of British Columbia (UBC), Vancouver, BC, Canada
shiyuliang@sdu.edu.cn, {han.yu, zqshen}@ntu.edu.sg

Abstract

Social insurance benefits qualification assessment is an important task to ensure that retirees enjoy their benefits according to the regulations. It also plays a key role in curbing social security frauds. In this paper, we report the deployment of the Intelligent Benefit Certification and Analysis (IBCA) platform, an AI-empowered platform for verifying the status of retirees to ensure proper dispursement of funds in Shandong province, China. Based on an improved Gated Recurrent Unit (GRU) neural network, IBCA aggregates missing value interpolation, temporal information, and global and local feature extraction to perform accurate retiree survival rate prediction. Based on the predicted results, a reliability assessment mechanism based on Variational Auto-Encoder (VAE) and Monte-Carlo Dropout (MC Dropout) is executed to perform reliability assessment. Deployed since November 2019, the IBCA platform has been adopted by 12 cities across the Shandong province, handling over 50 terabytes of data. It has empowered human resources and social services, civil affairs, and health care institutions to collaboratively provide high-quality public services. Under the IBCA platform, the efficiency of resources utilization as well as the accuracy of benefit qualification assessment have been significantly improved. It has helped Dareway Software Co. Ltd earn over RMB 50 million of revenue.

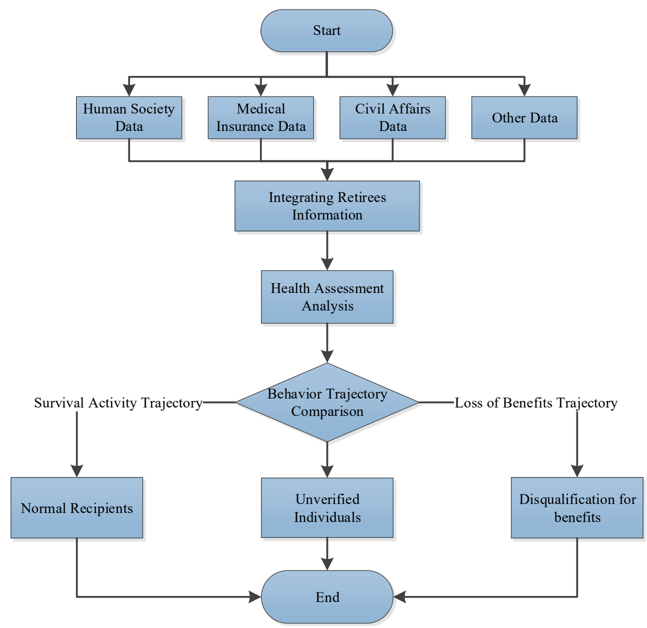

Figure 1: The process of social insurance benefits certification through data comparison.

Introduction

The process of qualifying applicants for social insurance benefits involves periodic verification by the pension agencies to confirm that pensioners continue to meet the eligibility criteria. In the context of China's increasingly aging population, the number of pensioners receiving these benefits has been consistently growing. The qualification certification for social insurance benefits is a crucial undertaking which ensures that retirees can access social security benefits in accordance with the law. This process is essential not only for retirees' well-being but also to safeguard the integrity of social security funds by preventing fraudulent and excessive pension claims.

*Corresponding author

The payment of pension insurance benefits across various government services involves multiple departments, including the Human Resources and Social Security Department and the Civil Affairs Department. The timely reporting of the demise of entitled recipients is of utmost importance, as any delay can lead to over payment of pensions and put pressure on the national pension funds. To ensure proper usage of the pension funds, government services across the board mandate regular qualification assessment and certification for benefit recipients. This proactive approach aims to reduce fraudulent claims, thereby ensuring the integrity of the social insurance system.

The prevailing method of social insurance qualification certification in China (Wang 2001) includes the following options:

1. **Face Recognition Self-Service Authentication**: This in-

volves authentication through a mobile app by utilizing facial recognition technology.

2. **Assisted Social Service Authentication**: To assist elderly and other special groups with limited mobility, a door-to-door certification service is available. Social security agents pay home visits to assist with the certification process.
3. **Data Comparison**: This step entails gathering retirees' information through various data comparison techniques.

At present, the third method is widely adopted in China for benefit qualification authentication. Figure 1 illustrates the authentication process through data comparison. By leveraging data from various governmental services, this approach captures the trajectory of retirees' survival-related activities through manual auditing and data analysis. This information helps determine the health and survival status of individuals, allowing for the identification of those eligible to receive benefits, those who have lost their benefit qualification, and those requiring recertification. This authentication method faces the following challenges:

1. **Accuracy and Completeness of Data**: The qualification certification process for insurance benefits involves the collection of a substantial volume of personal information and relevant documents. Unfortunately, due to the scattered and inconsistent nature of this information, ensuring the accuracy and completeness of the data is a significant challenge.
2. **Accuracy and Validity of Certification**: There are concerns regarding the precision and validity of current certification outcomes. A notable disparity exists between these certification results and the ground truth, suggesting room for improvement.
3. **Efficiency and Promptness of Certification**: The conventional approach to insurance benefit qualification predominantly relies on manual review and processing. Unfortunately, this method is not only time-consuming and labor-intensive but is also susceptible to errors.

To address the aforementioned concerns, we propose the Intelligent Benefit Certification and Analysis (IBCA) platform. At the core of this platform lies a robust artificial intelligence (AI) engine that integrates a novel survival prediction model, referred to as the Missing Value Interpolation based Temporal Survival Prediction (MVI-TSP) approach. This model combines diverse elements, including missing value interpolation, temporal information integration, comprehensive global and local feature extraction, and the utilization of a Gated Recurrent Unit (GRU) neural network (Cho et al. 2014), all of which synergistically enhance the precision of survival prediction for retirees. Furthermore, the IBCA platform innovatively incorporates a reliability assessment mechanism, leveraging Variational Auto-Encoder (VAE) (Caterini, Doucet, and Sejdinovic 2018) and Monte-Carlo Dropout (MC Dropout) (Sadr, Zhu, and Hu 2023). This advanced mechanism computes uncertainty scores for both the missing value interpolation process and the model prediction results, increasing reliability on two fronts. The platform provides in-depth analytical insights, promoting transparent decision support (Li et al. 2022).

Jointly developed by Dareway Software Co. Ltd.[1] and the Centre for AI Research (C-FAIR) jointly established by Shandong University, China and Nanyang Technological University, Singapore, the platform has been adopted by 12 municipalities in Shandong province, including regional population centres such as Jinan, Zibo, Yantai and Weifang. It has helped various governmental bodies, including big data bureaus, human resources, social departments, and other public services to collaboratively ensure the proper dispursement of pension funds. It has handled massive data volumes exceeding 50 terabytes since its deployment in November 2019. Under the IBCA platform, resource utilization has been improved whilst misuse of resources has been significantly reduced. The platform has yielded substantial economic gains, with a cumulative contract revenue reaching RMB 52.637 million for Dareway.

Application Description

In this section, we discuss in detail the system design of the IBCA platform. The system architecture, as shown in Figure 2, consists of four main modules:

1. **Data Preprocessing**: The primary role of this module is first to gather essential data from diverse sources such as government information resource sharing platforms, encompassing personal details, medical insurance, and civil affairs information. This encompasses fields within healthcare, health, and well-being. Subsequently, the collected data undergo rigorous cleaning, integration, and conversion before being formatted as input data for relevant intelligent analysis models.
2. **Modeling and Analysis**: This module predominantly comprises survival prediction model. The model draws from data shared by various governmental resources, considering factors such as age, gender, and disease type among retirees to estimate individual or group health status and survival probabilities. Survival prediction serves

[1]http://www.dareway.com.cn/

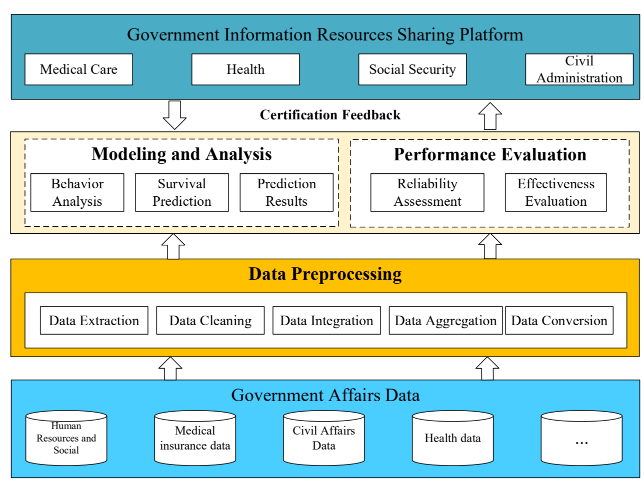

Figure 2: The architecture of the IBCA platform.

as the foundational element for optimizing subsequent assessments, and the accuracy of this prediction significantly influences the validity of downstream processes. The proposed AI model achieves the required accuracy for our objectives.

3. **Performance Evaluation**: The role of this module is pivotal in assessing the accuracy, reliability, and adaptability of the prediction model. This assessment is essential for ensuring the effectiveness and safety of AI applications. Performance evaluation gauges the reliability of the model output, ensuring consistent results across diverse datasets and samples, rather than excelling solely on specific datasets. Moreover, it aids in determining the model's applicability across various populations and environments, highlighting its robust generalization capabilities.
4. **Certification Feedback**: This module provides assessments and feedback on the benefit qualification results back into the business systems of each domain, facilitated by the government information resource sharing platform. The primary purpose of this feedback loop is to evaluate the quality and security of government services, thereby fostering continuous improvement and enhancement within government organizations.

Use of AI Technology

The application of AI techniques in the `IBCA` platform focuses on two main aspects: 1) survival prediction, and 2) reliability assessment. In this section, we describe these two parts in detail.

Survival Prediction

The key feature of the platform is the precise prediction of the health and survival status of retirees. This prediction, in turn, furnishes more accurate data for verifying the survival status of individuals and facilitating the accurate disbursement of pension benefits. Currently, survival prediction analysis faces the following challenges:

1. **Missing Data**: Missing data are mainly due to malfunctioning data measurement equipment, software glitches, or inadequate data recording. They result in poor quality of results from data analysis. Therefore, the development of a method for interpolating missing data is important.
2. **Temporal Information**: When dealing with data as a time series, not only does the survival status of retirees evolve over time, but the impact of the missing information also evolves temporally. Hence, accounting for the influence of time information is important when interpolating missing values.
3. **Hidden Information Extraction**: Many existing methods for survival prediction focus on analyzing standard factors such as age, gender and disease type, often neglecting the influence of secondary factors derived from these standard factors.

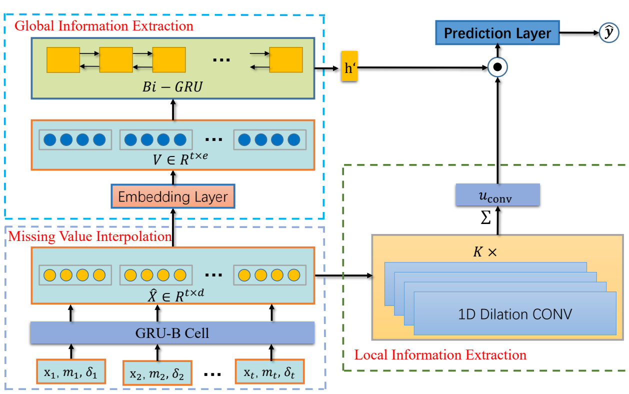

Figure 3: The architecture of `MVI-TSP`.

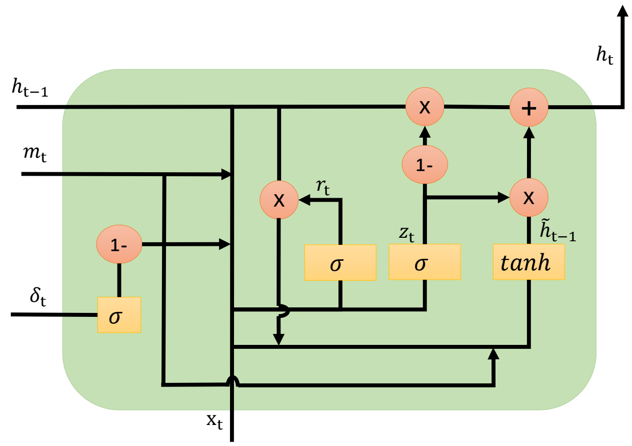

Figure 4: The architecture of GRU-B.

To overcome these challenges, we have introduced a pioneering survival prediction model, the Missing Value Interpolation-based Temporal Survival Prediction (`MVI-TSP`) model, as a core component of the `IBCA` platform, as illustrated in Figure 3. The model includes missing value interpolation, comprehensive global and local information acquisition, and precise survival prediction. To tackle the missing data issue, this paper proposes an enhanced internal structure of the GRU unit, referred to as GRU-B (as depicted in Figure 4). Building upon the bidirectional GRU architecture, GRU-B introduces a mask matrix and a time decay function to capture potential correlation relationships within the data, enabling the crucial functionality of missing value interpolation. Additionally, `MVI-TSP` leverages a one-dimensional convolutional neural network (Yu and Koltun 2016) and a bidirectional recurrent neural network (Schuster and Paliwal 1997) for information feature extraction from both local and global perspectives. This collective approach enables the model to effectively capture the survival patterns of retirees, substantially enhancing the accuracy of survival prediction. More details about `MVI-TSP` can be found in (Li et al. 2022).

Reliability Assessment

After missing data interpolation, existing methods often overlook the significance of incorporating uncertainty or confidence measures into the survival prediction modeling. Such oversight can lead to challenges in modeling and a decline in performance (Gal and Ghahramani 2016). Throughout the modeling process, it becomes paramount for survival prediction to be guided by uncertainty scores. Otherwise, overly confident prediction models can introduce issues for the organization (Tan et al. 2019; Jun et al. 2021). Therefore, the inclusion of confidence-driven survival prediction models stands as a pivotal requirement within the realm of benefits certification analysis. The `IBCA` platform prioritizes the optimization of confidence assessment, addressing both the uncertainty introduced during the missing value interpolation phase and the level of confidence associated with the resulting survival prediction outcomes.

Handling Missing Data Interpolation Uncertainty. For the task of missing data interpolation, we adopt the Variational Recurrent Neural Network (VRNN) approach (Chung et al. 2015; Bertugli et al. 2021). This VRNN framework introduces valuable uncertainty measures following the missing value interpolation, which serves as complementary information, enabling the model to predict patient survival more accurately. Specifically, the VRNN employs a Variational Auto-Encoder (VAE) (Lopez et al. 2018; Muthukumaran, Hariharanath, and Haridasan 2023) within each time step. This entire process can be summarized in four key steps: 1) Prioritization (*Priori*): in this step, potential variables are prioritized based on the state of the Recurrent Neural Network (h_{t-1}). 2) *Inference*: the encoder is utilized to approximate the posterior distribution of the observation variable (x) and the state variable, yielding the potential random variable (z). 3) *Generation*: using the decoder, an approximate distribution of x is generated based on the potential random variable (z). 4) *Update of State Variable*: the state variable (h) is updated using the Recurrent Neural Network (RNN). Implementation details of this process are as follows.

Priori. The priors of the latent random variables follow this distribution:

$$\begin{aligned} & z_t \sim N(\mu_{0,t}, diag(\sigma_{0,t}^2)), \\ & \text{where } [\mu_{0,t}, \sigma_{0,t}] = F^{prior}(h_{t-1}). \end{aligned} \tag{1}$$

Among them, $\mu_{0,t}$ and $\sigma_{0,t}$ denote the parameters of the conditional prior distribution, and F^{prior} is a function that takes the hidden state variables of the previous time step as input.

Inference. We leverage the encoder to learn a network that approximates the posterior distribution of a latent random variable, and the function F^{enc} to estimate the mean and log variance, conditioned on x_t and h_{t-1}:

$$\begin{aligned} & z_t \mid x_t \sim N\left(\mu_{z,t}, diag\left(\sigma_{z,t}^2\right)\right), \\ & \text{where } [\mu_{z,t}, \sigma_{z,t}] = F^{enc}\left(F^X\left(x_t\right), h_{t-1}\right). \end{aligned} \tag{2}$$

Among them, $\mu_{z,t}$ and $\sigma_{z,t}$ represent the parameters to approximate the posterior, and F^X is the feature extractor of x_t. We utilize the reparameterization method (Rezende, Mohamed, and Wierstra 2014) to perform network gradient descent. Specifically, we sample $e \sim N(0,1)$, and then make $z = \mu_{z,t} + \sigma_{z,t} * e$, where $*$ denotes element-wise multiplication.

Generation. The generative distribution is conditional on h_{t-1} in addition to z_t, so that:

$$\begin{aligned} & x_t^{'} \mid z_t \sim N\left(\mu_{x,t}, diag\left(\sigma_{x,t}^2\right)\right), \\ & \text{where } [\mu_{x,t}, \sigma_{x,t}] = F^{dec}\left(F^Z\left(z_t\right), h_{t-1}\right) \end{aligned} \tag{3}$$

Among them, $\mu_{x,t}$ and $\sigma_{x,t}$ represent the parameters of the generated distribution, F^Z is the feature extractor of z_t, and F^{dec} is the decoder. In this way, we obtain the mean value and the corresponding variance of the variables generated by the VRNN at time step t. In addition, RNNs update their hidden states using a recursive equation:

$$h_t = F^{RNN}(F^X(x_t), F^Z(z_t), h_{t-1}) \tag{4}$$

where F^{RNN} is the network of RNN units. From this equation, we find that h_t is a function of x_t and z_t. Therefore, Eq. (1) and Eq. (3) define the distributions $p(z_t|x_{<t}, z_{<t})$ and $p(x_t'|z_{\leq t}, x_{<t})$, respectively. The joint distribution of the generative model is:

$$p\left(x_{\leq T}, z_{\leq T}\right) = \prod_{t=1}^{T} p\left(x_t \mid z_{\leq t}, x_{<t}\right) p\left(z_t \mid x_{<t}, z_{<t}\right). \tag{5}$$

Similarly, the distribution of the inference model is:

$$q\left(z_{\leq T}, x_{\leq T}\right) = \prod_{t=1}^{T} q\left(z_t \mid x_{\leq t}, z_{<t}\right). \tag{6}$$

The `IBCA` platform leverages the Variational Recurrent Neural Network (VRNN) for modeling both the mean and variance of missing values during the missing value interpolation process. In this endeavor, the platform introduces the utilization of a mask matrix (M) and temporal information (δ) to compute the uncertainty post-interpolation. Specifically, the mask matrix (M) functions as an indicator, revealing whether a value is missing at a specific time point, while the temporal information (δ) indicates the duration for which a value remains missing at that given moment in time. Furthermore, the model output encompasses latent random variables which significantly enhance the modeling of intricate time series data (Chung et al. 2015). Similar to the previously mentioned GRU-B unit, we first define a time decay rate γ_t by using a sigmoid function, i.e.

$$\gamma_t = 1 - sigmoid\left(\delta_t\right). \tag{7}$$

Let $\sigma_{x,t}$ be the variance computed by the generative network. We use $\sigma_{x,t}$ to compute the uncertainty score for the imputed value. If the value is observed, the uncertainty is 0; otherwise, the larger the variance, the larger the uncertainty. Then,

$$u_t = sigmoid((1 - M) * \sigma_{x,t}) \tag{8}$$

where u_t denotes the uncertain fraction of the interpolated data.

Prediction Result Uncertainty. We define W as the weight matrix of the single hidden layer network, b as the bias, x as the input, and y as the output. Thus, we have:

$$y = sigmoid(Wx + b). \tag{9}$$

We set a random variable vector d that obeys Bernoulli distribution (Gal and Ghahramani 2016) to multiply by the weight matrix to perform dropout. Specifically, $d \sim Bernoulli(p), p \in [0,1]$, and the input is mapped to the output as:

$$y = sigmoid((dW)x + b). \tag{10}$$

The Bernoulli function is to randomly generate a vector of 0 and 1 with probability p. According to variational inference (Gal and Ghahramani 2016), the predicted distribution is:

$$q\left(y^{*} \mid x^{*}\right)=\int p\left(y^{*} \mid x^{*}, \omega\right) q(\omega)\, d\omega \tag{11}$$

where y^* represents the output. x^* represents the input. ω represents the parameter weight. We sample T Bernoulli and weight distributions and use Monte-Carlo to estimate the first moment of the prediction function as:

$$\mathbb{E}_{q(y^*|x^*)}\left(y^{*}\right) \approx \frac{1}{T} \sum_{t=1}^{T} \widehat{y}^{*}\left(x^{*}, W_{1}^{t}, \ldots, W_{L}^{t}\right). \tag{12}$$

Among them, W_1^t,...,W_L^t represents the matrix composed of T groups of vectors sampled from Bernoulli distribution. Similarly, we can obtain the variance of the model as:

$$\begin{aligned} &Var_{q(y^*|x^*)}\left(y^{*}\right) \approx \tau^{-1} I_{D} \\ &+\frac{1}{T} \sum_{t=1}^{T} \widehat{y}^{*}\left(x^{*}, W_{1}^{t}, \ldots, W_{L}^{t}\right)^{\top} \widehat{y}^{*}\left(x^{*}, W_{1}^{t}, \ldots, W_{L}^{t}\right) \\ &-\mathbb{E}_{q(y^*|x^*)}\left(y^{*}\right)^{\top} \mathbb{E}_{q(y^*|x^*)}\left(y^{*}\right) \end{aligned} \tag{13}$$

where τ is set to 1, and I_D is the D-dimensional identity matrix. In practice, this is equivalent to performing T random forward passes in the network and finding the variance. In our model, we use MC dropout (Sadr, Zhu, and Hu 2023) to estimate the predicted labels and uncertainty scores of our model as follows:

$$\begin{aligned} &Prediction : \mathbb{E}_{p(\widehat{y})}[\widehat{y}], \\ &Uncertainty : Var_{p(\widehat{y})}[\widehat{y}]. \end{aligned} \tag{14}$$

Application Development and Deployment

The IBCA platform was implemented using Java and JSP programming languages, developed by Dareway Software Co. Ltd in Jinan, Shandong Province, China. The platform uses a Hadoop infrastructure, enabling efficient storage of vast quantities of shared data from diverse governmental resources. In the development of the AI engine, we conducted a comprehensive evaluation of five established survival prediction analytical models, each serving a distinct purpose:

1. **Logistic Regression (LR)** (DeStefano 1990): This is a foundational algorithm in machine learning which uses linear regression normalized by a sigmoid function.
2. **RetainEX** (Kwon et al. 2019): This model introduces a time decay factor and attention mechanism, enhancing its utility for medical task prediction.
3. **AdaCare** (Ma et al. 2020): Utilizing dilated convolution with multiple scales, this model captures concealed insights into long and short-term historical data changes, employing GRU for prediction tasks.

	AUROC	AUPRC
LR	0.7145±0.0005	0.2302±0.0003
RetainEX	0.9665±0.0004	0.8640±0.0003
AdaCare	0.9671±0.0007	0.8476±0.0002
Diople	0.9689±0.0005	0.8697±0.0003
GRU-D	0.9772±0.0005	0.8748±0.0004
MVI-TSP	**0.9841±0.0006**	**0.9158±0.0011**

Table 1: Survival prediction offline test results.

4. **Diople** (Ma et al. 2017): This model harnesses a bidirectional recurrent neural network to memorize past and future visit information, integrating three attention mechanisms to quantify inter-visit relationships for predictive analysis.
5. **GRU-D** (Che et al. 2016): This model addresses missing data directly by amalgamating masking and time intervals within the GRU architecture.

The performance of these models was rigorously tested on a government dataset[2] from a city within Shandong Province, China. This dataset consists of a one-year temporal span across varying seasons. The five candidate methods were tasked with predicting the survival of individuals in a forthcoming time period. The results, as shown in Table 1, demonstrate that MVI-TSP has the highest AUROC (Area Under the Receiver Operating Characteristic Curve) and AUPRC (Area Under the Precision-Recall Curve) values of 0.9841 and 0.9158, respectively, based on the test dataset. In comparison to the best-performing alternative method (i.e., GRU-D), MVI-TSP exhibits improvements of 0.69% and 4.1% in AUROC and AUPRC values, respectively. Here we need to clarify that the main objective of this paper is to predict the survival status of treatment eligibles based on historical behavioral data, so as to derive the list of to-be-tested with high probability of death risk, and then determine the survival information of to-be-tested by manual verification. The comparisons depicted in Table 1 are intended to verify whether the proposed model is better able to tap into the population with high probability of death risk so that it can be applied to the developed platform. Consequently, MVI-TSP was selected for implementation in the AI engine of the IBCA platform.

Figure 5 shows the user interface of the IBCA certification process, encompassing various vital components: the certification analysis of the benefit population, trend analysis, multi-channel trajectory data, and the latest statistics on the certified population. Specifically: 1) Benefit Population Certification: This section provides a snapshot of the total number of certified individuals, the certification rate, and their corresponding changes over time. 2) Trend Analysis: The trend analysis segment offers insights into the certification status of individuals with varying trust levels, focusing on the preceding six months. 3) Multi-channel Trajectory Data: This segment showcases the utilization of trajec-

[2]The dataset is available upon request via email to the contact authors as the requesters will be asked to agree on certain terms of use.

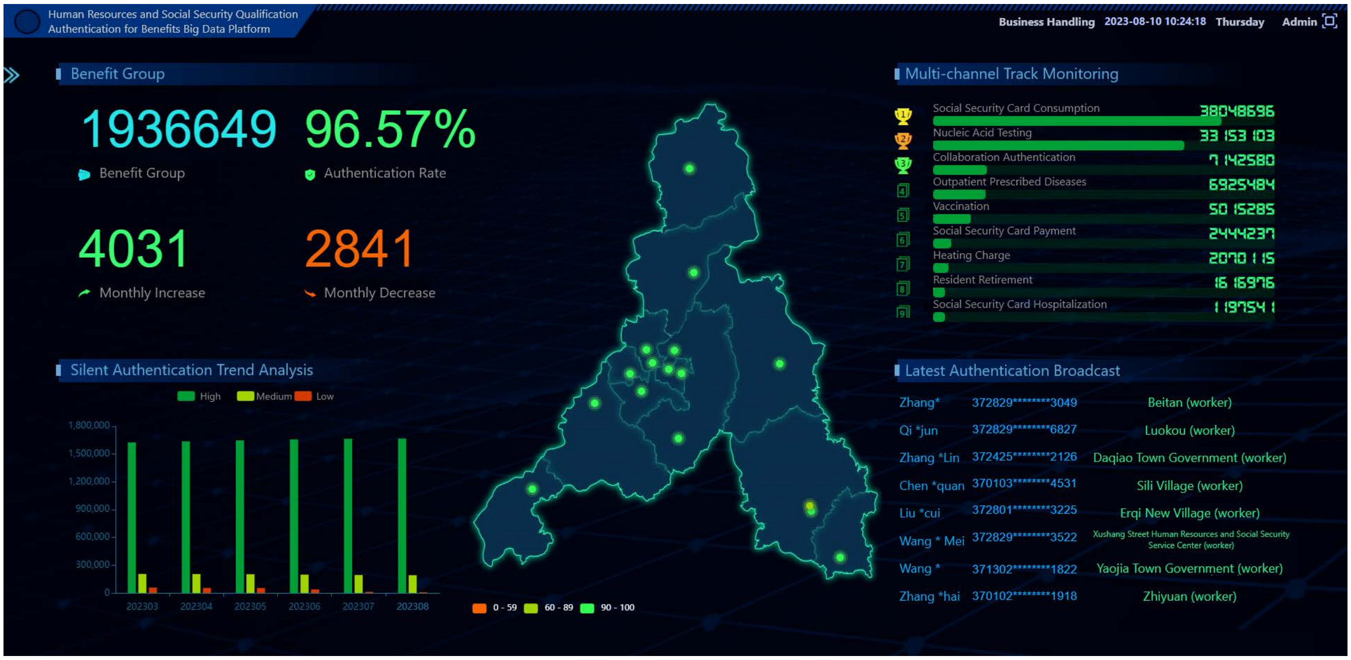

Figure 5: The IBCA platform interface for the overall certification process.

Figure 6: The IBCA platform interface for survival prediction and analysis results.

tory data across multiple channels, offering a comprehensive view of data access patterns on the platform. 4) Latest Certification Statistics: This section offers up-to-date information on certified individuals, providing a current snapshot of the certified population.

Figure 6 is a screenshot showing information and survival prediction results pertaining to trajectories within social insurance, health insurance, and other relevant domains. In addition, detailed analyses are conducted for each trajectory type, including the number of individuals certified at different trust levels, the corresponding percentages, and the validity periods for each trajectory. Additionally, this figure offers an insight into the incremental trends observed in recent months.

Figure 7: The IBCA platform interface for the qualified population, certification fit, and certification by institution.

Figure 7 focuses on the in-depth analysis of the benefits qualified population, authentication fit, and certifications based on the respective affiliations of the individuals. 1) Benefits Qualified Population: Analyzes different categories of certified individuals, such as employees, residents, and other certified personnel, providing a breakdown based on these categories. 2) Authentication Fit: Shows the number and percentage of individuals within different classes, as determined by the reliability assessment model. 3) Authentication by Institution: Offers a detailed analysis of the number of individuals certified under each administrative institution they are affiliated with, providing insights into the distribution of certifications across different institutions.

Note that the interfaces illustrated in Figures 5-7 are translated into English for the benefit of non-Chinese speaking readers. The actual deployed system interfaces are in Chinese for the target users. A video demonstration of the IBCA platform can be found on Youtube[3].

Application Use and Payoff

In this section, we discuss the real-world impact and pivotal role played by the IBCA platform. It has pioneered a novel, unobtrusive approach to benefits qualification certification, upending the traditional process. By harnessing aggregated multi-channel business data, it constructs a comprehensive database for benefits qualifications, leveraging big data technology to conduct analysis and certification. This approach offers real-time and effective insights into the survival status of retirees, facilitating prompt cessation of pension payments for deceased individuals. Concurrently, it significantly reduces the workload of government service personnel involved in business handling, leading to significant reductions in administrative costs.

Since its deployment in November 2019, the IBCA platform has gained widespread adoption in Shandong, making a significant impact within various governmental domains, including big data bureaus and human resources departments. This operation has resulted in a substantial data processing volume of 50 terabytes, empowering diverse sectors such as human resources and social services, civil affairs, health, and construction, among others.

The platform significantly enhances the accuracy of survival prediction for retirees through the comprehensive integration of multi-model AI technology. Among the 80,028 residents normally receiving benefits, 611,421 have had hospitalization records within the past 6 months, with 1,137 deaths in January. The preliminary analysis of the survival prediction model reveals that 780 of the 1,137 deceased individuals were accurately predicted by IBCA, alongside 78,891 survivors. The predicted survival count is 70,893, yielding a recall rate of 68.8% and an accuracy rate of 89.6%. For the remaining 531,393 individuals without hospitalization records, among the 679 who passed away in one month, the model predicted 509 accurately, with 530,714 survivors. The predicted survival count is 526,912, resulting in a recall rate of 75.0% and an accuracy rate of 99.3%.

In addition, the products stemming from the IBCA platform have been effectively marketed, utilizing strategies such as system upgrades, new product promotions, platform expansion, and the provision of comprehensive big data services. These efforts have yielded remarkable economic benefits, with cumulative revenue reaching RMB 52.637 million, a testament to the commercial viability and positive so-

[3]https://www.youtube.com/watch?v=DGy5FqqAyik

cietal impact of the IBCA platform.

Maintenance

The workflows, personnel and operational parameters in IBCA may change with time. Updates can be performed without impacting the AI engine due to the separation of concerns achieved through a modular system design. These AI algorithms have not required modification since their deployment in November 2019.

Lessons Learned During Deployment

The real-world deployment of the IBCA platform has highlighted the following key lessons:

1. The target users are mostly not information technology (IT) savvy with little knowledge of AI. The IBCA platform follows modular design approach to clearly separate the responsibilities of different stakeholders so that the required IT skills level is low. This, coupled with standardization through various information templates, has reduced the level of training required for new users of the platform. The detailed outputs of the AI engine are only presented to the high level decision-makers and technical team. This enables the technical team to provide more target support during deployment.
2. The IBCA platform was designed for efficiency at the expense of data privacy protection. Certain operations such as the sharing of digital assets amongst different government organizations do not follow best practices for data privacy protection. In subsequent development, we will adopt a react, resolve, reinvent approach to balance the need for efficiency with the need for ethical AI solutions (Tzachor et al. 2020). Beyond the initial "react" phase, we plan to incorporate privacy preserving machine learning paradigms such as federated learning (Yang et al. 2020; Kairouz et al. 2021) to enhance the protection of sensitive information. The aims is to transform the IBCA platform into an exemplary AI for social good solution (Tomašev et al. 2020).

Conclusions and Future Work

This paper discusses the use of AI technology to solve challenges encountered in benefits eligibility certification in Shandong Province, China. We have developed the IBCA platform to improve the survival status verification process for retirees and pension disbursements. IBCA leverages big data analysis to capture the trajectory of retirees' health related information and predict their health and survival statuses. Since its deployment in November 2019, the platform has been successfully adopted across 12 major cities in Shandong Province. It has been embraced by governmental departments, including big data bureaus, human resources and social services, among others. This widespread adoption empowers areas such as human resources, social services, civil affairs, health, construction, and disability services, providing substantial benefits to these sectors. The platform has also been a commercial success.

In subsequent work, we intend to explore the application of interpretable AI methods (Zeng et al. 2019; Theunissen and Browning 2022) to automatically generate explanations for AI predictions and recommendations (Yu et al. 2016, 2017). This effort aims to enhance understanding within governmental organizations, fostering acceptance and trust in AI (Yu et al. 2014a,b; Tanyel, Ayvaz, and Keserci 2023). Given the significant amount of personal information involved in the benefits certification process, e.g., ID numbers and medical records, data security and privacy protection are paramount. Hence, we will investigate the integration of privacy-preserving machine learning techniques, such as federated learning (Gao et al. 2019; Fu et al. 2022; Goebel et al. 2023), into the IBCA platform. This approach will enable collaboration among various governmental departments, while complying with privacy-protecting regulations such as the General Data Protection Regulation (Goddard and Michelle 2017).

Acknowledgments

This work is supported by the Key R&D Program of Shandong Province, China (2021CXGC010103); the National Natural Science Foundation of China (Grant No. 62376135); Joint SDU-NTU Centre for Artificial Intelligence Research (C-FAIR) (NSC-2019-011); the National Research Foundation Singapore and DSO National Laboratories under the AI Singapore Programme (AISG Award No: AISG2-RP-2020-019); and the RIE 2020 Advanced Manufacturing and Engineering (AME) Programmatic Fund (No. A20G8b0102), Singapore. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore.

References

Bertugli, A.; Calderara, S.; Coscia, P.; Ballan, L.; and Cucchiara, R. 2021. AC-VRNN: Attentive Conditional-VRNN for multi-future trajectory prediction. *Comput. Vis. Image Underst.*, 210: 103245.

Caterini, A. L.; Doucet, A.; and Sejdinovic, D. 2018. Hamiltonian Variational Auto-Encoder. In *NeurIPS*.

Che, Z.; Purushotham, S.; Cho, K.; Sontag, D. A.; and Liu, Y. 2016. Recurrent Neural Networks for Multivariate Time Series with Missing Values. *CoRR*, abs/1606.01865.

Cho, K.; van Merrienboer, B.; Gülçehre, Ç.; Bahdanau, D.; Bougares, F.; Schwenk, H.; and Bengio, Y. 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. In *EMNLP*, 1724–1734.

Chung, J.; Kastner, K.; Dinh, L.; Goel, K.; Courville, A. C.; and Bengio, Y. 2015. A Recurrent Latent Variable Model for Sequential Data. In *NeurIPS*, 2980–2988.

DeStefano, J. J. 1990. Logistic regression and the Boltzmann machine. In *IJCNN*, 199–204.

Fu, X.; Zhang, B.; Dong, Y.; Chen, C.; and Li, J. 2022. Federated Graph Machine Learning: A Survey of Concepts, Techniques, and Applications. *SIGKDD Explor.*, 24(2): 32–47.

Gal, Y.; and Ghahramani, Z. 2016. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. In *ICML*, volume 48, 1050–1059.

Gao, D.; Liu, Y.; Huang, A.; Ju, C.; Yu, H.; and Yang, Q. 2019. Privacy-preserving Heterogeneous Federated Transfer Learning. In *IEEE BigData*, 2552–2559. IEEE.

Goddard; and Michelle. 2017. The EU General Data Protection Regulation (GDPR): European regulation that has a global impact. *International Journal of Market Research*, 59(6): 703.

Goebel, R.; Yu, H.; Faltings, B.; Fan, L.; and Xiong, Z. 2023. *Trustworthy Federated Learning*, volume 13448. Springer, Cham.

Jun, E.; Mulyadi, A. W.; Choi, J.; and Suk, H. 2021. Uncertainty-Gated Stochastic Sequential Model for EHR Mortality Prediction. *IEEE Trans. Neural Networks Learn. Syst.*, 32(9): 4052–4062.

Kairouz, P.; et al. 2021. Advances and Open Problems in Federated Learning. *Foundations and Trends in Machine Learning*, 14(1-2): 1–210.

Kwon, B. C.; Choi, M.; Kim, J. T.; Choi, E.; Kim, Y. B.; Kwon, S.; Sun, J.; and Choo, J. 2019. RetainVis: Visual Analytics with Interpretable and Interactive Recurrent Neural Networks on Electronic Medical Records. *IEEE Trans. Vis. Comput. Graph.*, 25(1): 299–309.

Li, B.; Shi, Y.; Cheng, L.; Yan, Z.; Wang, X.; and Li, H. 2022. MTSSP: Missing Value Imputation in Multivariate Time Series for Survival Prediction. In *IJCNN*, 1–8.

Lopez, R.; Regier, J.; Jordan, M. I.; and Yosef, N. 2018. Information Constraints on Auto-Encoding Variational Bayes. In *NeurIPS*, 6117–6128.

Ma, F.; Chitta, R.; Zhou, J.; You, Q.; Sun, T.; and Gao, J. 2017. Dipole: Diagnosis Prediction in Healthcare via Attention-based Bidirectional Recurrent Neural Networks. In *KDD*, 1903–1911.

Ma, L.; Gao, J.; Wang, Y.; Zhang, C.; Wang, J.; Ruan, W.; Tang, W.; Gao, X.; and Ma, X. 2020. AdaCare: Explainable Clinical Health Status Representation Learning via Scale-Adaptive Feature Extraction and Recalibration. In *AAAI*, 825–832.

Muthukumaran, K.; Hariharanath, K.; and Haridasan, V. 2023. Feature Selection with Optimal Variational Auto Encoder for Financial Crisis Prediction. *Comput. Syst. Sci. Eng.*, 45(1): 887–901.

Rezende, D. J.; Mohamed, S.; and Wierstra, D. 2014. Stochastic Backpropagation and Approximate Inference in Deep Generative Models. In *ICML*, volume 32, 1278–1286.

Sadr, M. A. M.; Zhu, Y.; and Hu, P. 2023. An Anomaly Detection Method for Satellites Using Monte Carlo Dropout. *IEEE Trans. Aerosp. Electron. Syst.*, 59(2): 2044–2052.

Schuster, M.; and Paliwal, K. 1997. Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.*, 45: 2673–2681.

Tan, Q.; Ma, A. J.; Ye, M.; Yang, B.; Deng, H.; Wong, V. W.; Tse, Y.; Yip, T. C.; Wong, G. L.; Ching, J. Y.; Chan, F. K.; and Yuen, P. C. 2019. UA-CRNN: Uncertainty-Aware Convolutional Recurrent Neural Network for Mortality Risk Prediction. In *CIKM*, 109–118. ACM.

Tanyel, T.; Ayvaz, S.; and Keserci, B. 2023. Beyond Known Reality: Exploiting Counterfactual Explanations for Medical Research. *CoRR*, abs/2307.02131.

Theunissen, M.; and Browning, J. 2022. Putting explainable AI in context: institutional explanations for medical AI. *Ethics Inf. Technol.*, 24(2): 23.

Tomašev, N.; Cornebise, J.; Hutter, F.; Mohamed, S.; Picciariello, A.; Connelly, B.; Belgrave, D. C. M.; Ezer, D.; van der Haert, F. C.; Mugisha, F.; Abila, G.; Arai, H.; Almiraat, H.; Proskurnia, J.; Snyder, K.; Otake-Matsuura, M.; Othman, M.; Glasmachers, T.; de Wever, W.; Teh, Y. W.; Khan, M. E.; Winne, R. D.; Schaul, T.; and Clopath, C. 2020. AI for social good: Unlocking the opportunity for positive impact. *Nature Communications*, 11(2468): doi:10.1038/s41467–020–15871–z.

Tzachor, A.; Whittlestone, J.; Sundaram, L.; and hÉigeartaigh, S. O. 2020. Artificial intelligence in a crisis needs ethics with urgency. *Nature Machine Intelligence*, 2: 365–366.

Wang, F. 2001. The Identification of the Qualification of Enjoying the Unemployment Insurance in China. *Journal of Anshan Teachers College*, 3(4): 43–45.

Yang, Q.; Liu, Y.; Cheng, Y.; Kang, Y.; Chen, T.; and Yu, H., eds. 2020. *Federated Learning*. Springer, Cham.

Yu, F.; and Koltun, V. 2016. Multi-Scale Context Aggregation by Dilated Convolutions. In *ICLR*.

Yu, H.; Miao, C.; An, B.; Shen, Z.; and Leung, C. 2014a. Reputation-aware Task Allocation for Human Trustees. In *AAMAS*, 357–364.

Yu, H.; Miao, C.; Chen, Y.; Fauvel, S.; Li, X.; and Lesser, V. R. 2017. Algorithmic management for improving collective productivity in crowdsourcing. *Scientific Reports*, 7(12541).

Yu, H.; Miao, C.; Leung, C.; Chen, Y.; Fauvel, S.; Lesser, V. R.; and Yang, Q. 2016. Mitigating herding in hierarchical crowdsourcing networks. *Scientific Reports*, 6(4).

Yu, H.; Shen, Z.; Miao, C.; An, B.; and Leung, C. 2014b. Filtering trust opinions through reinforcement learning. *Decision Support Systems*, 66: 102–113.

Zeng, Z.; Miao, C.; Leung, C.; Shen, Z.; and Chin, J. J. 2019. Computing Argumentative Explanations in Bipolar Argumentation Frameworks. In *AAAI*, 10079–10080.