

## Article

# Estimation of Uncertainty for Technology Evaluation Factors via Bayesian Neural Networks

Juhyun Lee <sup>1</sup>, Sangsung Park <sup>2,\*</sup> and Junseok Lee <sup>3,\*</sup><sup>1</sup> Institute of Engineering Research, Korea University, Seoul 02841, Republic of Korea<sup>2</sup> Department of Data Science, Cheongju University, Chungbuk 28503, Republic of Korea<sup>3</sup> Machine Learning Big Data Institute, Korea University, Seoul 02841, Republic of Korea

\* Correspondence: hanyul@cju.ac.kr (S.P.); jxli12@korea.ac.kr (J.L.)

**Abstract:** In contemporary times, science-based technologies are needed for launching innovative products and services in the market. As technology-based management strategies are gaining importance, associated patents need to be comprehensively studied. Previous studies have proposed predictive models based on patent factors. However, technology-based management strategies can influence the growth and decline of firms. Thus, this study aims to estimate uncertainties of the factors that are frequently used in technology-based studies. Furthermore, the importance of the factors may fluctuate over time. Therefore, we propose a Bayesian neural network model based on Flipout and four research hypotheses to evaluate the validity of our method. The proposed method not only estimates the uncertainties of the factors, but also predicts the future value of technologies. Our contribution is to (i) provide a tractable Bayesian neural network applicable to big data, (ii) discover factors that affect the value of technology, and (iii) present empirical evidence for the timeliness and objectivity of technology evaluation. In our experiments, 3781 healthcare-related cases of patents were used, and we found that the proposed hypotheses were all statistically significant. Therefore, we believe that reliable and stable technology-based management strategies can be established through our method.

**Keywords:** technology evaluation; uncertainty; Bayesian neural network; patent; evaluation factor

**MSC:** 62H12; 62P99; 62F15; 68U35



**Citation:** Lee, J.; Park, S.; Lee, J. Estimation of Uncertainty for Technology Evaluation Factors via Bayesian Neural Networks. *Axioms* **2023**, *12*, 145. <https://doi.org/10.3390/axioms12020145>

Academic Editor: Oscar Humberto Montiel Ross

Received: 27 December 2022

Revised: 27 January 2023

Accepted: 28 January 2023

Published: 31 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Kaufmann and Tödting (2001) emphasized that firms should develop scientific technologies to launch innovative products in the market [1]. At present, firms are actively reflecting technologies in their management plans. They employ technologies to (i) pinpoint their strengths [2–4], (ii) compare them with those of competitors [5–7], and (iii) develop new business models [8,9]. Furthermore, firms create intellectual property (IP) from their management activities based on their technology, which comprises the IP research and development (IP-R&D) of firms.

Generally, IP-R&D employs patents, which are typical IPs. Firms can claim their rights to specific technologies, such as patents. Thus, firms proactively use patents for prior arts, prediction of emerging technology, and technology evaluations. These patents explain (i) how novel technology is, (ii) how advanced technology is compared to prior art, and (iii) how technology contributes to industrial development. Patents can also explain the features of technologies such as originality, marketability, and scope of rights using various factors.

Patents include citations, family patents, and claims. Citations refer to the number of cited papers of prior art or cited information by other patents. The originality of technology can be measured based on the citations of the patent [10–12]. As firms want

their technologies to be widely used, they register patent families in a number of nations. Therefore, patent families can be used as factors to estimate the marketability of the technology [13]. Claims refer to the scope of rights of the technology. The higher the number of claims is, the wider the scope of rights to protect the patent [14,15].

IP-R&D using patent factors has been studied in various forms. Firms have several IP-R&D strategies that can be used to develop innovative products and services. First, firms can reduce the cost of R&D through technology trade or transfer [16,17]. Second, the prediction of vacant technology can help the discovery of new business models for firms [18]. Third, firms can predict emerging technology to invest in their limited budgets efficiently and intensively [19]. Finally, technology evaluation can be used in the planning of various IP-R&D activities such as the prediction of technology transfer, vacant technology, and emerging technology [20,21].

IP-R&D can also help predict the future of technology through various factors of patents. These advantages act as a sufficient condition for industry, academia, and even the government to select IP-R&D. However, we need to consider the uncertainty of patent factors to establish more complete IP-R&D strategies. This is because the basis of IP-R&D is trust in factors that explain the features of technology.

The purpose of this study is to estimate the uncertainty of various patent factors used in IP-R&D. In particular, we focus on technology evaluation for the following reasons.

- Technology evaluation that can predict technology excellence is a bridge that connects various types of IP-R&D and firms.
- The technology value may fluctuate depending on the time of evaluation.
- The uncertainty of technology value can be actively reflected in the management and investment of firms.

Thus, we estimate the uncertainty of factors used in technology evaluation in order to (i) help firms make the right decisions, (ii) support timely R&D, and (iii) contribute towards minimizing losses. To achieve this, this study proposes a Bayesian neural network (BNN) model for technology evaluation. The proposed methodology not only estimates the uncertainty of evaluation factors but also predicts the patent value.

The contributions of the proposed methodology are as follows:

- Jun (2022) and Uhm and Jun (2022) pointed out the intractable limitations of the Bayesian approach as the volume of big data gradually increased [22,23]. To improve these limitations, we applied Flipout as a tractable and appropriate approach to big data. Flipout helped BNN learn each layer independently.
- Lee and Park (2022) emphasized that identifying factors influencing IP-R&D can prevent the absence of validity in patent analysis [24]. Thus, we measured the influence on the technology value for each evaluation factor. To this end, the difference between the mean and variance of the value according to the factor was statistically tested.
- Choi et al. (2023) argued that for sustainable growth in an uncertain business environment, it is necessary to cope with the rapidly changing flow of technology [25]. Therefore, we measured the uncertainty of technology evaluation over time. In addition, we verified the trend and presented empirical evidence for the timeliness and objectivity of technology evaluation.

The rest of the paper is structured as follows. Sections 2 and 3 explain the previous studies on technology evaluation and the background of the Bayesian neural network, respectively. Section 4 proposes a method for estimating technology evaluation factors. Section 5 describes an experiment to demonstrate the applicability of our method. Finally, Section 6 discusses several limitations of our study, and Section 7 presents suggestions for future studies.

## 2. Related Works

In the past, technology evaluations were conducted based on the opinions of technology experts. However, as patent data increase and computing power improves, a large

number of studies on social network analysis or the use of machine learning and deep learning have been conducted.

Previously, technology evaluation has been conducted via the expert-based Delphi method [26,27]. However, such methodologies potentially produced biased evaluations depending on the opinion of individual experts. Furthermore, they did not reflect various factors included in patents and could not predict the future technology value. To overcome these limitations, Galbraith et al. (2006) proposed a combination of technology evaluation factors and an expert-based method [28]. Furthermore, Akoka and Comyn-Wattiau (2017) pointed out the risk of evaluation of new technologies and stressed the need for a methodology that can reflect this risk [29]. Sa et al. (2022) suggested a Delphi-based methodology, noting that the opinions of experts and academics are important to capturing and evaluating the value of technology [30].

Patents cite prior art to protect rights and technological progress. Owing to this characteristic, several researchers analyzed patents using a form of a social network. Kim et al. (2015) used connected patent information in a network to evaluate the excellence of technology [31]. Choi et al. (2015) and Kumari et al. (2021) predicted potential opportunities for commercialization by discovering hidden knowledge of technology through a combination of machine learning and social networks [32,33]. Lai et al. (2023) proposed a methodology using main path analysis to predict the potential impact on the development of technology [34]. They extracted the contributions of patents and linked them together to trace the technology trajectory. As a result, they were able to discover the trends in which technology has developed, and among them, they found technologies with high value. In contrast to the Delphi-based method, they proposed a data-based model, thereby contributing towards ensuring the objectivity of evaluation.

Yang et al. (2012) stated that the patent value was determined by factors such as citations and claims [35]. They aimed to secure the objectivity of evaluation by estimating the royalty rate of technology. Trappey et al. (2012) mentioned the need for selecting only the main information among various evaluation factors [36]. Through the proposed method, they can predict the future value of technology using a deep neural network, which was trained with the selected main factors. Trappey et al. (2019) and Ko et al. (2019) noted that patents were transferable assets having innate economic and technological value [37,38]. In this regard, Chung et al. (2020), Lee et al. (2022), and Huang et al. (2022) proposed an evaluation model based on machine learning and deep learning to extract various values of patents [39–41]. Nonetheless, their methods cannot estimate the uncertainty of evaluations.

Table 1 presents the comparison results of previous studies on technology evaluation based on features. Expert-based approaches lack the objectivity of evaluation. Social network analysis-based approaches could mitigate the limitations of expert-based approaches. However, they could not predict the value of future technology. Machine learning- and deep learning-based approaches can overcome the drawbacks of existing approaches. However, they cannot measure the uncertainty of technology evaluations. In this paper, our contributions are (i) to ensure the objectivity of technology evaluation, (ii) to predict the future technology value, and (iii) to measure the uncertainty to building effective IP-R&D.

**Table 1.** Comparison of previous studies on technology evaluation.

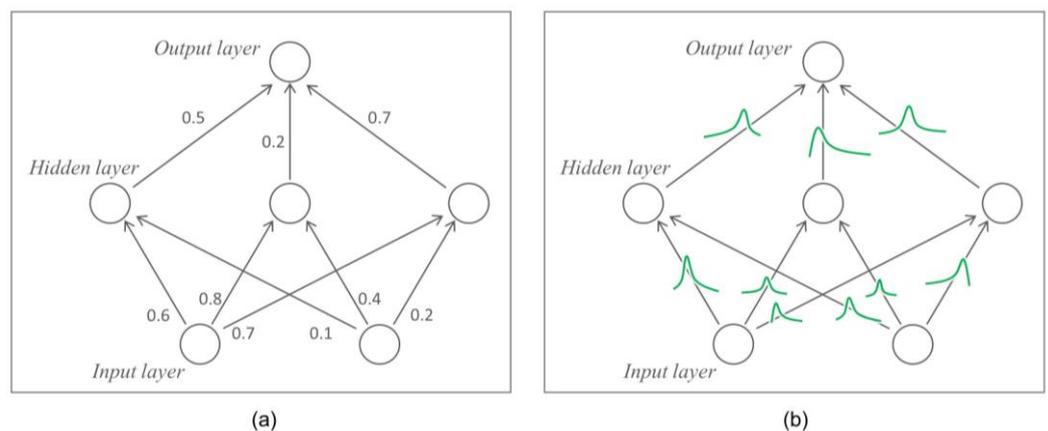
Approaches	Related Works	Objectivity	Prediction	Uncertainty
Expert-based	[26–30]	–	–	–
SNA-Based <sup>1</sup>	[31–34]	✓	–	–
ML and DL Based <sup>2</sup>	[35–41]	✓	✓	–
BNN-based	Our method	✓	✓	✓

<sup>1</sup> Social Network Analysis-based approaches. <sup>2</sup> Machine and Deep Learning-based approaches.

### 3. Background

An artificial neural network is a predictive model that simulates the actions of neurons, which are found in the human brain. A neural network-based model broadly consists of three layers. Features of the observed values enter the input layer. Then, data inputted to an input layer are converted to predicted values after passing through hidden and output layers. Let us assume that  $N$  pieces of data that enter a neural network are  $\mathcal{D} \sim \{x_i, y_i\}_{i=1}^N$ . Then, the neural network learns a fixed weight  $W$  that connects layers from  $\mathcal{D}$ .

In the past, neural networks have suffered from (i) incomplete learning in nonlinear separable space, (ii) vanishing gradient as the network deepens, and (iii) overfitting of the training dataset. Nonetheless, a neural network is currently one of the most popular predictive models and is used in Lecun’s (1988) back propagation (BP) algorithm [42], Nair and Hinton’s (2010) rectified linear units (ReLU) [43], and Srivastava et al. (2014)’s dropout [44]. The deep neural network (DNN) architecture has attracted attention not only because of its use of BP, ReLU, and dropout but also its data quality and computer power development. DNN is built by having deep hidden layers that are present between the input and output layers. Furthermore, the application range of DNN has gradually become wider as it uses various layers such as convolutional or recurrent layers. Figure 1a shows the simple DNN architecture. DNN predicts a data label with the fixed weight  $W$  that is learned using a training dataset.



**Figure 1.** Comparison of deep neural network and Bayesian neural network. (a) a deep neural network is a constant with a fixed weight connecting each layer; (b) in a Bayesian neural network, the weight connecting each layer is a probability distribution.

Here, let us assume that neural network  $f$  used in an autonomous driving system determines “stop” and “go” based on the front image. In this example, the neural network learns as it receives a large number of images as input data  $x$ . A modern DNN can learn without overfitting the training dataset while eliminating the vanishing gradient for nonlinear separable spaces using BP, ReLU, and dropout. Let us assume that random images are inputted into the learned DNN. Then,  $f(x, W)$ , which is a predictive value of the DNN, is either “go” and “stop”. That is, a fixed weight-based DNN has a tendency to be overconfident in the training dataset.

DNN is used in areas that are closely associated with human life such as autonomous driving and healthcare. However, given the potential for accidents due to overconfident DNNs of the training dataset, an alternative is now needed. BNN, which handles a weight  $W$  that connects layers as random variables is one of the alternatives of DNN [45–48]. Figure 1b shows the BNN architecture. The most crucial difference between DNN and BNN is the viewpoint of weight  $W$ . In contrast with DNN that uses a fixed weight, weight  $W$  is assumed as a random variable of  $q_\theta$  in BNN. Thus, BNN estimates a probability distribution  $q_\theta$  with Bayes’ theorem. However, because the operation of posterior in Bayes’ theorem is intractable, it is difficult for BNN training to converge. To overcome this, various tractable

mechanisms were proposed such as Wan et al. (2013)'s DropConnect [49] including dropout and Kingma et al. (2015)'s reparameterization trick [50,51].

BNN predicts a sample label using various weights  $W$  that are extracted from  $q_\theta$ . That is, BNN cannot measure the uncertainty of the prediction. Thus, BNN helps the analyzer to make the right decision by increasing the uncertainty of prediction if an unseen case emerges in a training dataset. Wen et al. (2018) introduced Flipout, pointing out that the existing BNN is difficult to learn as the variance of the gradient is high because all training samples in the mini-batch share the same perturbation [52]. Let us assume that there is a weight  $W = \bar{W} + \Delta W$  that follows  $q_\theta$  where  $\bar{W}$  and  $\Delta W$  refer to mean weights and stochastic perturbation, respectively. Let us also assume  $\Delta \hat{W} \sim q_\theta$ . Then, Flipout ensures the decreasing variance of gradient estimates. Under the training sample  $x$ , one entry  $\mathcal{G}_x$  of stochastic gradient  $\nabla_\theta \mathcal{L}(y, f(x, \bar{W}, \Delta W))$  can be expressed in Equation (1):

$$\frac{\partial}{\partial \theta_i} \mathcal{L}(y, f(x, \bar{W}, \Delta W)), \tag{1}$$

where  $\mathcal{G}_x$  is a random variable that depends on  $x$  and  $\Delta W$ .

Let us assume that  $\mathcal{G}_B$  is a gradient averaged over a size of mini-batch  $M$ . Then, by using the law of total variance,  $\mathcal{G}_B$  can be decomposed into the variance of the exact mini-batch gradients and the estimation variance for a fixed mini-batch (see Equations (2)–(4)).

$$\alpha = \text{Var}_x(\mathbb{E}_{\Delta W}[\mathcal{G}_x|x]) + \mathbb{E}_x[\text{Var}_{\Delta W}(\mathcal{G}_x|x)], \tag{2}$$

$$\beta = \mathbb{E}_{x,x',\Delta \hat{W}}[\text{Cov}_{\Delta W}(\mathcal{G}_x, \mathcal{G}_{x'}|x, x', \Delta \hat{W})], \tag{3}$$

In Equation (2),  $\alpha$  refers to the variance of the gradients on individual samples. In Equation (3),  $\beta$  refers to the covariance of the estimation of stochastic perturbation  $\Delta W$ . As Flipout aims to remove perturbation that is shared in the mini-batch samples,  $\beta$  is removed.

$$\gamma = \mathbb{E}_{x,x'}[\text{Cov}_{\Delta \hat{W}}(\mathbb{E}_{\Delta W}[\mathcal{G}_x|x, \Delta \hat{W}], \mathbb{E}_{\Delta W}[\mathcal{G}_{x'}|x', \Delta \hat{W}]|x, x')], \tag{4}$$

$\gamma$  in Equation (4) refers to the term of covariance in  $\Delta \hat{W}$ . Thus, Flipout trains each layer independently by decreasing  $\alpha$  and  $\gamma$ .

#### 4. Proposed Method

This study focuses on estimating the uncertainty of factors that are used to evaluate technologies. Let us assume that factors  $x_i$  and value  $y_i$  of  $N$  technologies follow data distribution  $\mathcal{D} \sim \{x_i, y_i\}_{i=1}^N$ . Let us also assume that the value of technology that is classified with the binary class is  $y_i = \{0, 1\}$ . When the value of technology is high,  $y_i$  is 1. Then, BNN  $f$  evaluates the technology through input  $x$  and weights  $W$ . Here, if probability  $P(f(x, W) = k)$  is larger than the threshold  $c$ , the sample is classified to  $k$  and vice versa ( $k$  is 0 or 1). The estimator of uncertainty about input  $x$  whose true label is  $k$  is presented in Equation (5):

$$u_k = P(f(x, W|y = k) \neq k), \tag{5}$$

where  $f(x, W|y = k)$  refers to the predicted label of the sample whose true label is  $k$ .

This study assumes four research hypotheses. First, the distribution of evaluation factors will differ depending on the technological value. This hypothesis implies that factors can classify technologies clearly according to their value. Second, the BNN's layer will follow a specific distribution. We assume a layer's distribution when learning a BNN through the proposed method. If the learning of BNN is adequate, the layer would be approximated to the assumed distribution. Third, the performance of BNN will be better than other classifiers. This study estimates the uncertainty of evaluation factors through BNN. Thus, a BNN whose performance is better than other predictive models would increase the reliability of estimation. Finally, the uncertainty of technology whose value is evaluated as excellent will be higher for more recent technology is more recent. The

adequate evaluation of technology requires timeliness and objectivity more than immediacy. Furthermore, evaluation factors may be affected by the time flow. Thus, we assume that the uncertainty of technology, which has been evaluated as excellent in recent years, would be high.

**Hypothesis 1.** *The quantitative factors of patents have different distributions depending on the technology’s value.*

The proposed methodology aims to predict technology’s value using quantitative factors and estimate the uncertainty of the factors. Generally, factors such as “number of forward citations” and “number of family patents” are used to evaluate technologies.

The average of the  $p$ -th factor  $x_{i,y=k}^{(p)}$  for the  $i$ -th sample whose label is  $k$  is presented in Equation (6):

$$\hat{\mu}_{p,y=k} = \frac{1}{N_k} \sum_{i=1}^N x_{i,y=k}^{(p)} \tag{6}$$

where  $N_k$  refers to the sample size where  $y$  is  $k$ .

Equation (7) presents the variance of each factor depending on the technology’s value.

$$\hat{\sigma}_{p,y=k}^2 = \frac{1}{N_k - 1} \sum_{i=1}^N \left( x_{i,y=k}^{(p)} - \hat{\mu}_{p,y=k} \right)^2 \tag{7}$$

Thus, Hypothesis 1 means that the difference between  $\hat{\mu}_{p,y=k}$  and  $\hat{\sigma}_{p,y=k}^2$  according to  $y$  is statistically significant. Thus, the null hypotheses we assume are as follows:

$$H_0 : \hat{\mu}_{p,y=0} \geq \hat{\mu}_{p,y=1} \tag{8}$$

$$H_0 : \hat{\sigma}_{p,y=0}^2 = \hat{\sigma}_{p,y=1}^2 \tag{9}$$

where the equality and inequality signs of the alternative hypotheses are defined as the opposite to those of the null hypotheses.

Equation (8) is rejected if the factor of a patent whose technology’s value is high is larger than that whose technology’s value is low. Equation (9) means there is no difference in the deviation of the distribution depending on the technology’s value by comparing the deviation of the factor.

**Hypothesis 2.** *The layers of the Bayesian neural network follow a specific probability distribution.*

The proposed methodology predicts the technology’s value using a BNN. This study assumes that the weight  $W$  of the layer in the BNN follows the standard normal distribution. Let us assume that random variable  $Z$  conditioned on  $Z$ :  $Z \sim N(0, 1)$ . Then, Hypothesis 2 assumes Equation (10).

$$H_0 : W \approx Z, \tag{10}$$

where the alternative hypothesis for Hypothesis 2 is that weight  $W$  does not fall within the standard normal distribution.

When Equation (10) cannot be rejected, the mean and variance of  $W$  will be approximated to 0 and 1, respectively. Their skewness and kurtosis are also close to 0.

**Hypothesis 3.** *The Bayesian neural network has higher prediction performance than other models.*

Hypothesis 3 assumes that the performance of BNN would be higher than that of other classifiers. BNN can measure a loss of samples  $T$  times through the layer distribution. Thus, the loss of BNN is calculated by Equation (11).

$$\bar{\mathcal{L}} = \mathbb{E}_{(x,y) \sim \mathcal{D}, W \sim q_{\theta}} \left[ \frac{1}{T} \sum_{t=1}^T \mathcal{L}(f(x, W), y)^{(t)} \right], \tag{11}$$

where  $q_\theta$  and  $\mathcal{L}(f(x, W), y)^{(t)}$  refer to the distribution of  $W$  and the  $t$ -th measured loss of the sample.

Here, let  $\tilde{\mathcal{L}}$  be the loss of a predictive model other than BNN. Then, the null hypothesis for Hypothesis 3 is Equation (12).

$$H_0 : \tilde{\mathcal{L}} \leq \bar{\mathcal{L}}, \tag{12}$$

where the alternative hypothesis for Hypothesis 3 is that  $\tilde{\mathcal{L}}$  is less than  $\bar{\mathcal{L}}$ .

By rejecting Equation (12), we can ensure that the reliability of the estimation of uncertainty and the BNN performance is better than other predictive models.

**Hypothesis 4.** *The uncertainty in highly valued technology increases over time.*

Technology value is determined by marketability, originality, and scope of rights. However, these factors may be evaluated differently depending on the time when technology development is completed. For example, the number of forward citations tends to be smaller for patents that are more recently registered. Furthermore, technologies that were developed in the past may attract attention a decade later. Thus, we assume that the uncertainty of technology evaluation would fluctuate depending on the timeliness and objectivity of the technology. Then, the null hypothesis for Hypothesis 4 is given by Equation (13).

$$H_0 : \text{No trend exists in uncertainty of technology evaluation}, \tag{13}$$

According to Hypothesis 4, the uncertainty of technology value that is evaluated more recently is higher than those that are not recent. In general, existing IP-R&D did not consider when the technology value is evaluated. However, the technology evaluation of recently registered patents would fluctuate more. Thus, it is essential to establish an IP-R&D strategy according to the time of technology evaluation.

Our study aims to estimate the uncertainty when evaluating a technology using factors that contain technology features. BNN has an advantage in that it can predict technology value several times for the same sample as it learns a distribution of weights. Using this, we estimate the uncertainty of the factors. Let us assume  $\mathcal{G}_B$ , the gradient averaged over a mini-batch  $M$ , as the random variable obtained in the learning process. When  $B$  is  $\{x_m\}_{m=1}^M$ ,  $\mathcal{G}_B$  is given by Equation (14).

$$\mathcal{G}_B = \frac{1}{M} \sum_{m=1}^M \mathcal{G}(x_m, \Delta W_m), \tag{14}$$

where  $\mathcal{G}(x, \Delta W)$  denotes the gradient under the perturbation  $\Delta W$  for Sample  $x$  [52].

$$\text{Var}(\mathcal{G}_B) = \frac{1}{M} \alpha + \frac{M-1}{M} \gamma, \tag{15}$$

Flipout helps the layer distribution of each factor to be independent of each other. To achieve this, the proposed method aims to reduce the expected variance estimation  $\alpha$  of individual data points and covariance  $\gamma$  of layer distribution. Equation (15) presents the variance of  $\mathcal{G}_B$  that is calculated through Flipout in the proposed method.

### 5. Experimental Results

The purpose of the experiment is to study the applicability of the proposed method. We collected a patent dataset consisting of 3781 healthcare cases. In this section, we explain (i) the preparation process for the experiment, (ii) the statistical test on the research hypotheses, and (iii) the feature ablation trial.

### 5.1. Experimental Setup

We collected 3781 patents for healthcare, which were filed in the United States Patent and Trademark Office (USPTO). Data-based healthcare has been applied in a wide range of areas in recent years [53]. In particular, incorrect analysis in healthcare can lead highly risky such as the prediction of oxygen saturation [54] or clinical interpretation of somatic mutations in cancer [55]. Thus, the evaluation of healthcare technologies must be objective and rigorous.

Table 2 presents the evaluation factors extracted from the collected patents. The factors from *num\_app* to *num\_ipc* are factors that explain the features of the patents as well as the predictors of the predictive model. The *value* in the table indicates the target variable of the patent value. The *value* refers to the technology value provided in the patent database, which is classified into High (or 1) and Low (or 0). Thus, we designed a binary-class classification model that predicts a target variable with patent predictors.

**Table 2.** Variables used in the proposed model.

Factors	Description
<i>num_app</i>	Number of applicants who own patent rights
<i>num_bkw</i>	Number of backward citations
<i>num_clm</i>	Number of claims
<i>num_cpc</i>	Number of Cooperative Patent Classification (CPC) codes
<i>num_doc</i>	Number of cited non-patent documents
<i>num_famN</i>	Number of countries for family patents
<i>num_famP</i>	Number of family patents
<i>num_frw</i>	Number of forward citations
<i>num_inv</i>	Number of Inventors
<i>num_ipc</i>	Number of International Patent Classification (IPC) codes
Value	High (or 1) if the value of the patent is high, and Low (or 0) otherwise

Table 3 presents the data splitting results to train the BNN. The test dataset comprises 30% of 3781 cases. The rest of the data were split into training and validation datasets in the proportion 8:2. The BNN trained the training dataset and employed the validation dataset to prevent overfitting. The performance of the model whose learning was completed was measured using the test dataset.

**Table 3.** Results of splitting the data to train the model.

Dataset	High	Low	Total	Ratio
Raw dataset	2335	1446	3781	100%
Training dataset	1307	809	2116	56%
Validation dataset	327	203	530	14%
Test dataset	701	434	1135	30%

The BNN architecture used in the experiment is as follows: The number of neurons in the input layer is 10, and the number of neurons in the two hidden layers are 6 and 4. The ReLU was used as the activation function in the layer. The model was learned to minimize a negative log loss. In the learning, 256 mini-batch sizes and 50 epochs were used. The Kullback–Leibler divergence was employed to approximate the layer distribution. Figure A1 in Appendix A shows the detailed architecture of BNN.

### 5.2. Statistical Test for Research Hypothesis

In this study, four research hypotheses were proposed in order to prove the validity of our method. We prove the proposed hypotheses using statistical tests. The statistical tests were conducted sequentially from Hypotheses 1 to 4. All statistical hypotheses were tested at the 0.05 significance level.

### 5.2.1. Statistical Test for Hypothesis 1

Hypothesis 1 assumes that the evaluation factors of patents would have a statistically significant impact on the determination of technology value. Table 4 presents the statistical test results according to the factors. For example, if we calculate the average and standard deviation of the number of backward citations (*num\_bkw*) for patents whose value is high, they are 79.099 and 259.065. In contrast, the average and standard deviation of *num\_bkw* of patents whose *value* is low are 34.435 ( $\hat{\mu}_{p,y=0}$ ) and 86.937 ( $\hat{\sigma}_{p,y=0}$ ).

**Table 4.** Results of statistical test for Hypothesis 1.

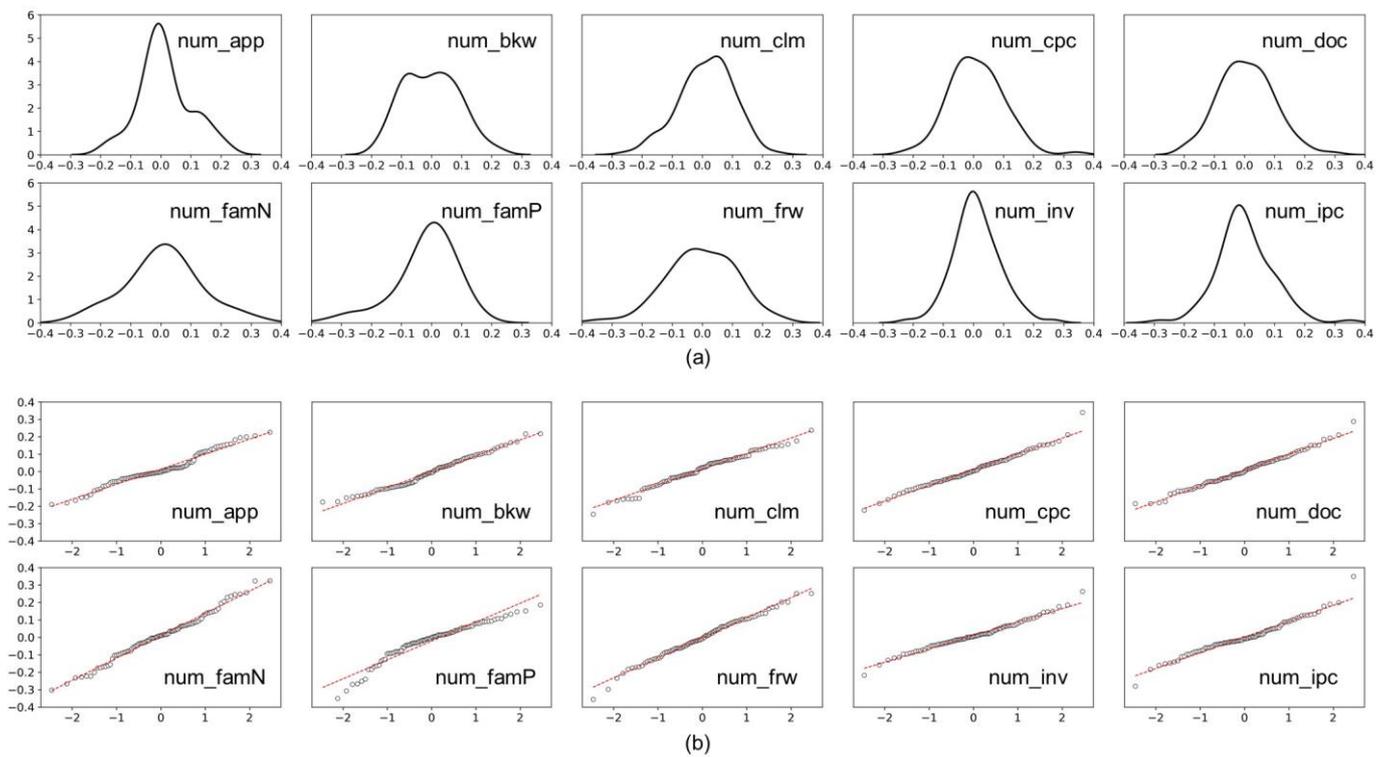
Variables	$\hat{\mu}_{p,y=0}$	$\hat{\mu}_{p,y=1}$	$\hat{\sigma}_{p,y=0}$	$\hat{\sigma}_{p,y=1}$	Levene *		T-Test *		Wilcoxon *	
					Statistics	p-Value	Statistics	p-Value	Statistics	p-Value
num_app	1.156	1.196	0.549	0.671	4.093	0.043	−1.931	0.027	−1.157	0.124
num_bkw	34.435	79.099	86.937	259.065	49.834	<0.001	−6.337	<0.001	−13.678	<0.001
num_clm	17.839	21.694	9.391	13.818	55.854	<0.001	−9.352	<0.001	−8.824	<0.001
num_cpc	8.686	13.376	7.676	16.573	154.033	<0.001	−10.107	<0.001	−6.720	<0.001
num_doc	7.737	18.119	32.677	52.314	52.232	<0.001	−6.770	<0.001	−11.455	<0.001
num_famN	2.652	5.535	2.277	4.022	354.661	<0.001	−24.888	<0.001	−25.482	<0.001
num_famP	14.008	87.576	47.859	287.619	131.161	<0.001	−9.641	<0.001	−26.605	<0.001
num_frw	14.728	74.661	24.087	115.006	398.873	<0.001	−19.546	<0.001	−33.123	<0.001
num_inv	2.215	2.094	1.931	2.269	34.145	<0.001	1.690	> 0.500	4.631	>0.500
num_ipc	4.060	4.416	3.231	5.785	49.939	<0.001	−2.146	0.016	6.442	>0.500

\* ‘Levene’ denotes the *p*-value of Levene’s test for homogeneity of variance. ‘T-test’ denotes the *p*-value of the T-test for homogeneity of average. ‘Wilcoxon’ denotes the *p*-value of the Wilcoxon Rank-Sum test, a nonparametric method for the T-test.

In the table, ‘Levene’ refers to the *p*-value of Levene’s test to test the equal variance of the factors according to the *value*. ‘T-test’ and ‘Wilcoxon’ refer to the *p*-values of the parametric and nonparametric methods to test whether the averages of the factors are equal according to the *value*. At the significance level of 0.05, the *num\_bkw* of patents whose value is high is more than that of patents whose value is low. Thus, *num\_bkw* can be viewed as a statistically adequate factor that can differentiate the technology’s value of the patents. The statistical test result of Hypothesis 1 showed that factors other than *num\_app*, *num\_inv*, and *num\_ipc* had a significant impact on determining the technology’s value of the patents.

### 5.2.2. Statistical Test for Hypothesis 2

Hypothesis 2 assumes that after training the BNN, the layer distribution is approximated to a normal distribution. Through this hypothesis, we can prove the adequacy of the training. Figure 2a shows the joint probability density distribution of the layer of each factor. Figure 2b shows the quantile–quantile plot (Q–Q plot) of the distribution of each layer. The results showed that most distributions of the factors were approximated to a normal distribution.



**Figure 2.** Visualization of the weight distribution of a trained Bayesian neural network. (a) Joint probability density distribution for the layers of each factor; (b) Q–Q plot of the distribution of each factor layer and standard normal distribution.

Table 5 presents the statistical test results of Hypothesis 2. We measured the average, standard deviation, skewness, and kurtosis of each layer’s distribution. The average and standard deviation of the standard normal distribution are 0 and 1, and their skewness and kurtosis are both 0. The comparison results of the basic statistics showed that the skewness and kurtosis of *num\_famP* were significantly different in the standard normal distribution.

**Table 5.** Results of statistical test for Hypothesis 2.

Variables	Avg	Std	Skewness	Kurtosis	Shapiro-Wilk *	
					Statistics	<i>p</i> -Value
num_app	0.013	0.086	0.278	0.081	0.965	0.010
num_bkw	−0.002	0.090	0.208	−0.676	0.980	0.133
num_clm	0.013	0.089	−0.333	0.032	0.985	0.316
num_cpc	0.010	0.090	0.389	0.985	0.986	0.369
num_doc	0.007	0.090	0.262	0.160	0.990	0.688
num_famN	0.009	0.126	0.045	0.112	0.988	0.519
num_famP	−0.022	0.112	−1.410	3.056	0.905	0.000
num_frw	−0.003	0.113	−0.286	0.212	0.991	0.722
num_inv	0.010	0.076	0.238	0.980	0.987	0.416
num_ipc	0.001	0.090	0.411	1.849	0.970	0.024

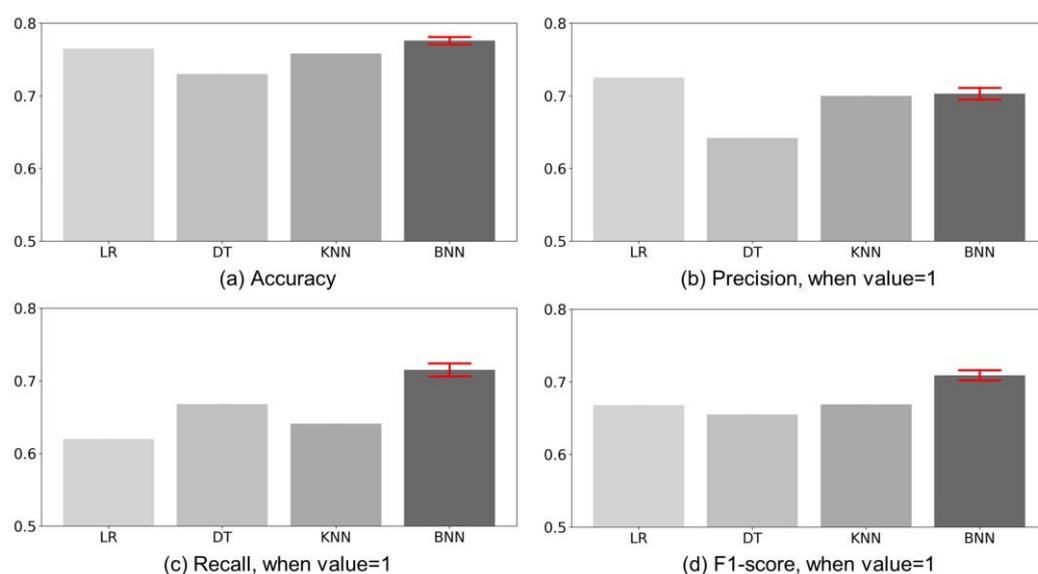
\* ‘Statistics’ and ‘*p*-value’ denote the statistics and *p*-value, respectively, of the Shapiro–Wilk test for normality.

The Shapiro–Wilk test of the technology evaluation factors revealed that *num\_app*, *num\_famP*, and *num\_ipc* rejected the null hypothesis at the significance level of 0.05 [56]. Thus, the distribution of the layer other than that of the above three factors can be viewed as following a normal distribution.

### 5.2.3. Statistical Test for Hypothesis 3

Hypothesis 3 assumes that the performance of the BNN would be higher than that of other predictive models. This ensures that the BNN not only estimates the uncertainty of the evaluation but also enables accurate prediction. The experiments were conducted to compare the performance of BNN and other predictive models. The comparison models were logistic regression (LR), decision tree (DT), and k-nearest neighbors (KNN). The Gini index was used to split a decision tree, and the number of neighbors used in KNN was five.

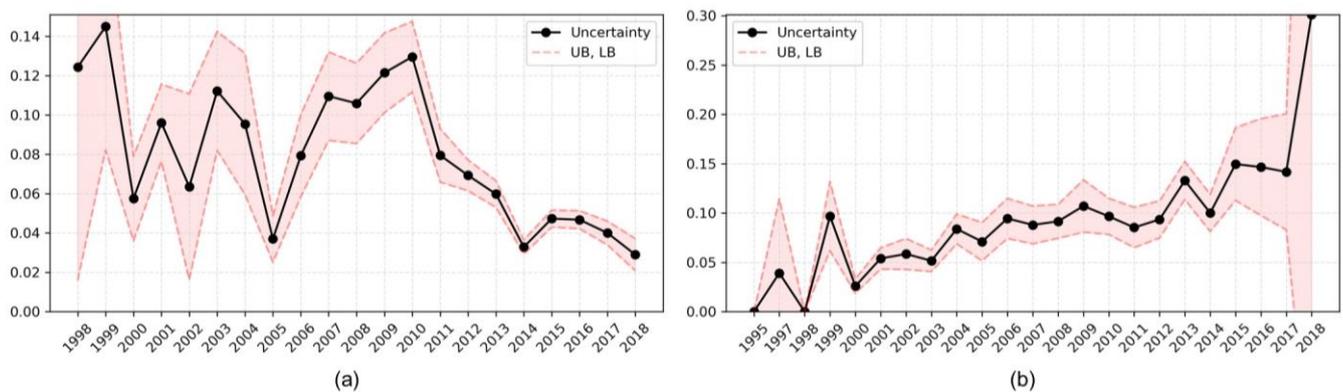
Figure 3 shows the comparison charts of (a) Accuracy, (b) Precision, (c) Recall, and (d) F1-score, which are indicators that measure the performance of the models. BNN calculated the prediction interval using the average and deviation, which was measured ten times ( $T = 10$ ) iteratively for the test dataset. In Accuracy, BNNs perform similarly to LRs. However, other performance of LR shows that the model's predictions are biased. The F1-score is an indicator showing the performance corrected for the bias of prediction. BNN's F1-score is higher than other classifiers. Specifically, BNN ([0.702, 0.716]), KNN (0.669), LR (0.668), and DT (0.655) are high in that order. The performance of BNN is up to 109.31% higher than other classifiers. Moreover, the range of variation is very small. In other words, BNN is (i) faster than prediction through expert evaluation, (ii) less costly, and (iii) stable prediction is possible. Thus, BNN can be seen as an adequate predictive model to estimate the uncertainty of the evaluation.



**Figure 3.** Comparison of predictive performance by model. (a) Measuring the accuracy of the models; (b) Precision when technology value is high (value = 1); (c) Recall when technology value is high (value = 1); (d) F1-score when technology value is high (value = 1).

### 5.2.4. Statistical Test for Hypothesis 4

Figure 4a shows the visualized results of the uncertainty according to the patent application year of patents whose value is low. BNN measured the uncertainty of patents whose technology value was low in the past very highly and vice versa. Figure 4b shows the uncertainty based on the application year of patents whose value is high. BNN measured the uncertainty of patents whose technology value was high in recent years very highly. Thus, researchers should be careful not to be overconfident regarding patents whose value is high among the recently invented technologies.



**Figure 4.** Visualization of trends in uncertainty over time. (a) Results for technologies evaluated as ‘Low’ in actual value; (b) Results for technologies evaluated as ‘High’ in actual value.

Table 6 presents the statistical test results of the uncertainty’s trend according to the patent application year. When the value was low, the results tended to decrease statistically significantly in the Spearman and Kendall rank correlation test. In contrast, when the value was high, the results tended to increase in the statistical test. Thus, the uncertainty of patents whose value was high gradually increased over time at the significance level of 0.05.

**Table 6.** Results of statistical test for Hypothesis 4.

Statistical Test	Spearman *		Kendall *	
	Statistics	<i>p</i> -Value	Statistics	<i>p</i> -Value
$u_0$	−0.594	0.005	−0.448	0.004
$u_1$	0.872	<0.001	0.737	<0.001

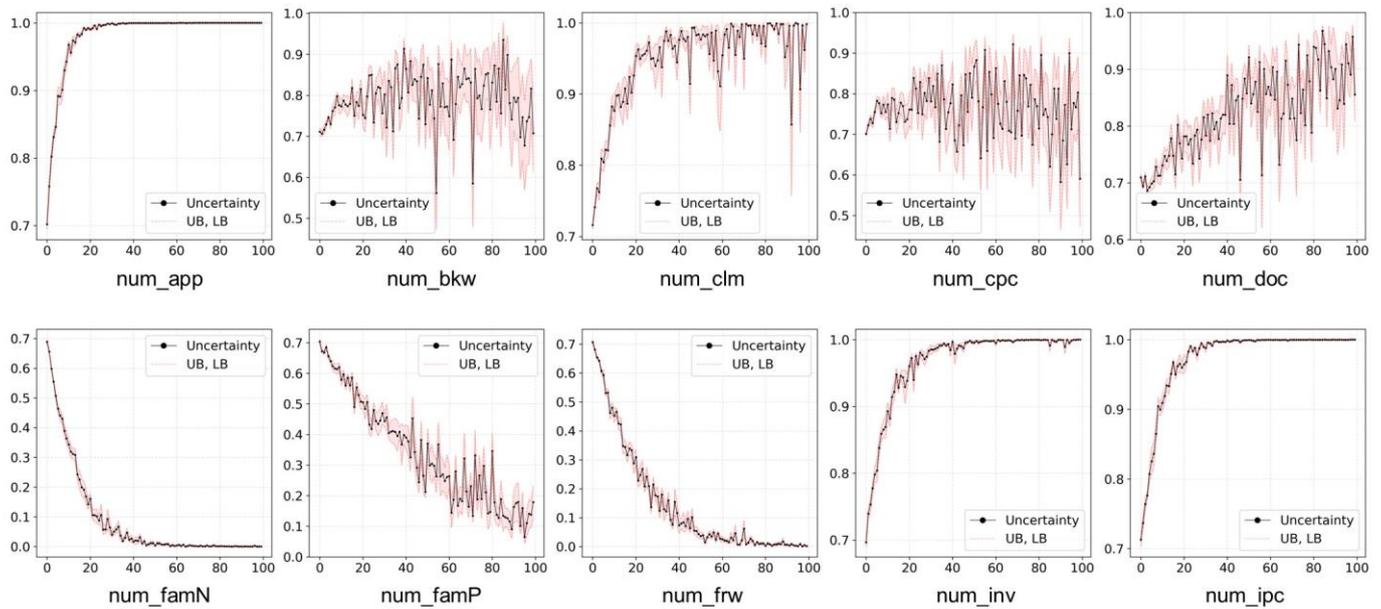
\* ‘Statistics’ and ‘*p*-value’ denotes the statistics and *p*-value, respectively, of Shapiro’s test for normality.

The experimental results using the patent data of healthcare exhibited that most factors had a significant impact on the evaluation of technology. Furthermore, the BNN’s weight, which learned patent data, followed a normal distribution, and its performance was better than that of other classifiers. The experiments also showed that the uncertainty of the technology’s value of recently registered patents was higher than that of the patents registered in the past.

### 5.3. Factor Ablation Trial

The purpose of the ablation study in deep learning is to observe the change in performance by removing a particular factor or layer in the network. This study aimed to measure the change in uncertainty by entering a random value to the factor in the BNN through the factor ablation trial. First, the operation of BNN for the rest of the factors except for the *p*-th factor is ignored. Then, the change in  $u_k$  is observed iteratively while increasing the value of the *p*-th factor from 0 to 100.

Figure 5 shows the results of the factor ablation trial. In the figure, the black line indicates the uncertainty of evaluation according to the factor. The red line denotes the upper bound (UB) and lower bound (LB) of fluctuating uncertainty. Through a factor ablation trial, we measured the degree to which the uncertainty of evaluation changes as the value of the factor increases.



**Figure 5.** Results of a factor ablation trial. The black line is the average of the uncertainty, and the red part represents the upper and lower bounds of the uncertainty.

Variables *num\_app*, *num\_clm*, *num\_doc*, *num\_inv*, and *num\_ipc* showed an increasing trend of uncertainty of technology evaluation as the factor increased. In contrast, the uncertainty of evaluation decreased as *num\_famN*, *num\_famP*, and *num\_frw* increased. In addition, *num\_bkw* and *num\_cpc* did not show a clear trend of increase or decrease in the uncertainty. Nevertheless, *num\_bkw* and *num\_cpc* had the lowest uncertainty around 60 and 90, respectively.

The results of the statistical test on the factor ablation trial are summarized in Table 7. We tested whether the trend of the uncertainty according to the change in the factors was increasing or decreasing using the Spearman and Kendall rank correlation test. At the significance level of 0.05, the results showed that factors *num\_frw*, *num\_famN*, and *num\_famP* contributed to decreasing the uncertainty of evaluation.

**Table 7.** Results of statistical tests on factor ablation trials.

Variables	Spearman *		Kendall *	
	Statistics	<i>p</i> -Value	Statistics	<i>p</i> -Value
<i>num_app</i>	0.981	<0.001	0.897	<0.001
<i>num_bkw</i>	0.143	0.077	0.116	0.044
<i>num_clm</i>	0.728	<0.001	0.588	<0.001
<i>num_cpc</i>	−0.047	0.321	−0.016	0.408
<i>num_doc</i>	0.788	<0.001	0.611	<0.001
<i>num_famN</i>	−0.973	<0.001	−0.887	<0.001
<i>num_famP</i>	−0.949	<0.001	−0.826	<0.001
<i>num_frw</i>	−0.972	<0.001	−0.878	<0.001
<i>num_inv</i>	0.916	<0.001	0.804	<0.001
<i>num_ipc</i>	0.972	<0.001	0.872	<0.001

\* ‘Statistics’ and ‘*p*-value’ means the statistics and *p*-value, respectively, of Shapiro’s test for normality.

The validation of the experiment is as follows:

- As a result, BNN had the highest F1-score, followed by KNN, LR, and DT. This means that the performance has improved by up to 109.31% compared to the previous one. Since the range of fluctuation of BNN performance is very small, the proposed method can make stable predictions.

- We found a trend in uncertainty over time. This means that when researchers build IP-R&D strategies, they must consider the period when the value of the technology was evaluated.
- Through the feature ablation trial, we found that the uncertainty of evaluation decreased as *num\_famN*, *num\_famP*, and *num\_frw* among the factors increased. In addition, we found that *num\_app*, *num\_clm*, *num\_doc*, *num\_inv*, and *num\_ipc* are proportional to the uncertainty of evaluation.

Through experiments, we were able to discover empirical evidence for the proposed research hypothesis. The strengths and weaknesses discovered along the way are discussed in the remaining sections.

## 6. Discussion

The importance of technology has been emphasized by various firms, universities, and research institutions in modern society. Specifically, firms search for their strengths or competitors' weaknesses through technologies. Patents are one of the important cornerstones to conducting IP-R&D because they exhibit the history of technological advances as well as predict the future vision.

Patents have been widely used to evaluate technological values in existing studies. They employed factors such as patent citations, claims, and family patents to train predictive models. However, as the importance of timeliness and objectivity is emphasized more than the immediacy of technology evaluation, it is necessary to have a study on methods to measure the uncertainty of the prediction.

The proposed method set four research hypotheses to estimate the uncertainty of technology evaluation factors. The first hypothesis was that factors would affect the technology value. The second hypothesis was that BNN learning would be adequate. The third hypothesis was that the BNN performance would be higher than that of other classifiers. The final hypothesis was that the more recent the evaluated value was, the higher the uncertainty was.

Empirical evidence obtained from statistical tests on the hypotheses is as follows:

- Technology value is determined by originality, marketability, and the scope of rights.
- The causal relationship of technology value with the number of applicants who own patent rights, the number of inventors, and the number of IPC codes should be carefully interpreted.
- Nonetheless, the evaluation of technology, which is higher in the number of countries for family patents, the number of forward citations, and the number of family patents, can be trusted more than that of other cases.

This study proposed a method to estimate the uncertainty of technology evaluation through Bayesian-based predictive models. Our experiments on the proposed method were conducted by collecting healthcare patents, which were likely to have a high risk of technology evaluation. We statistically tested four hypotheses using 3781 cases of data. The test results of the first hypothesis confirmed that there was a statistically significant difference in technology evaluation factors according to the patent value. More specifically, all factors except for the number of applicants who own patent rights, the number of inventors, and the number of IPC codes were fit to evaluate technologies.

Through the experiment results, we verified that the BNN layer, which learned healthcare-related patents, was approximated to the statistical distribution. The performance measurement results of technology evaluation exhibited that BNN had the highest F1-score ([0.702, 0.716]), followed by KNN (0.669), LR (0.668), and DT (0.655). We also verified that the uncertainty of the recently evaluated technology value was high. We discovered in the feature ablation trial that the larger the values of the number of countries for family patents (Spearman's statistics,  $-0.973$ ), the number of forward citations (Spearman's statistics,  $-0.972$ ), and the number of family patents (Spearman's statistics,  $-0.949$ ), the lower the uncertainty of technology evaluation. Thus, the results implied that these three factors were the important ones to determine the technology value.

This study estimated the uncertainty of factors used to evaluate the technology. IP-R&D is conducted in various ways, not only through technology evaluation but also technology transfer, vacant technology, and emerging technology. Thus, it is necessary to measure and compare the uncertainty of the tasks used in various IP-R&D strategies in future studies. Through this, we expect our study results to contribute to a wide range of IP-R&D activities.

## 7. Conclusions

Innovative technologies can be used by firms to develop new business models. Consequently, firms are interested in various IP-R&D activities such as the prediction of technology transfer, vacant technology, and emerging technology. Conventionally, IP-R&D was conducted through discussions between experts. However, this approach has a limitation that the result may differ depending on the experts' opinions. Much attention has been paid to the data-based approach by researchers because it can overcome this limitation.

Some researchers have focused on the fact that patents can be linked to citations and proposed a method using a social network. Other researchers have used machine learning or deep learning to predict the future technology value. However, the aforementioned approaches cannot estimate the uncertainty of modeling. Therefore, this study proposed a method that (i) ensures IP-R&D objectivity, (ii) predicts the future, and (iii) estimates the uncertainty.

This study has certain limitations:

- Patents are documents that describe the technology. However, we did not use patent texts.
- We predicted patent value through BNN. We also compared the performance with that of other classifiers. However, although our model was the best, it was not good enough to be applied to the real world.
- More recently, various layers have been proposed as deep learning advances. However, this was not considered.

Therefore, the future research directions we propose are as follows:

- In recent years, the study field of natural language processing has rapidly grown. Thus, it is necessary to consider the patent text when estimating technology value in future studies.
- Prediction performance needs to be further improved. Reliable prediction performance helps discover new technology opportunities beyond technology evaluation in various fields.
- We need to compare the estimation of uncertainty according to various layers in future studies. This is because the attention-based layer suitable for the patent document helps the model understand the technical terms of the patent [57].

We did not improve the prediction performance using various layers (e.g., convolutional layer) that is used in deep learning or patent texts. Nonetheless, we proposed various hypotheses to estimate the uncertainty of technology evaluation and verified the empirical evidence. Thus, we expect our contribution to be helpful for IP-R&D for the next industrial revolution.

**Author Contributions:** J.L. (Juhyun Lee) designed this research and conducted the experiment as described. S.P. collected the data set for the experiment. J.L. (Junseok Lee) analyzed the data to show the validity of this paper. In addition, all authors cooperated with each other in revising the paper. All authors have read and agreed to the published version of the manuscripts.

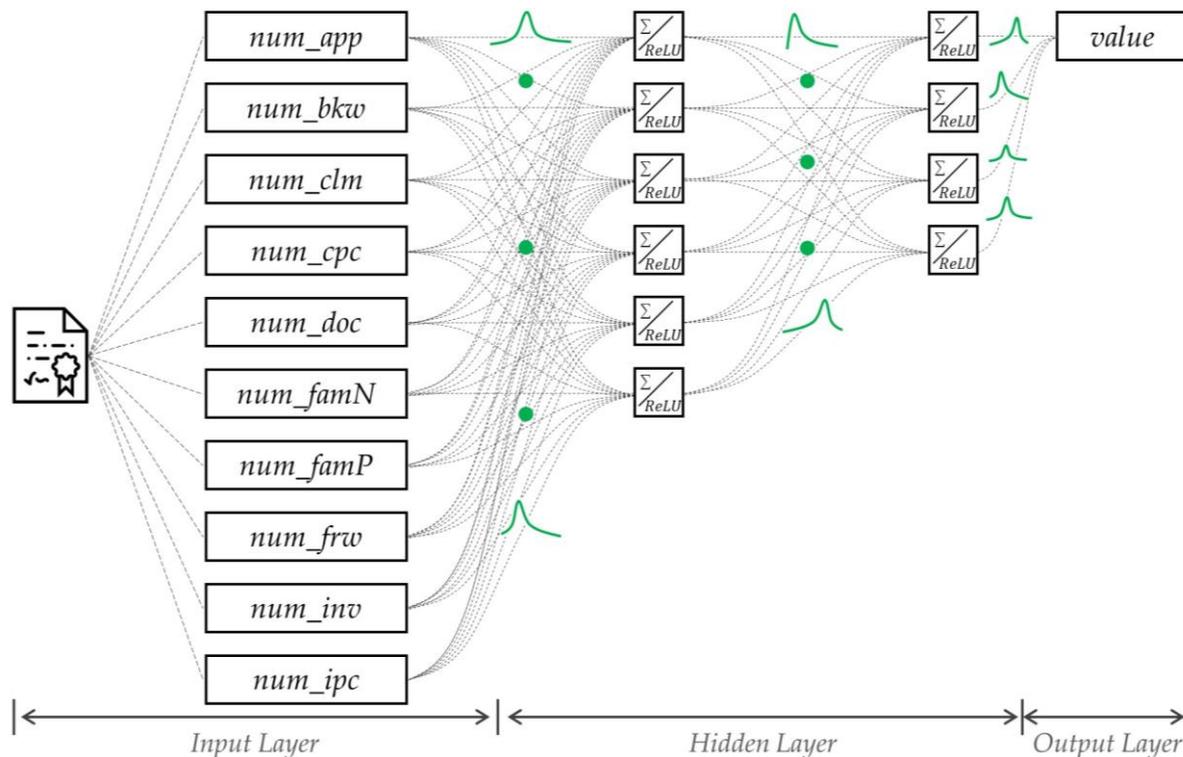
**Funding:** This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. NRF-2020R1A2C1005918). This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. NRF-2022R1I1A1A01069422).

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

Appendix A depicted the architecture of the Bayesian neural network model used in the experiment. We used Python 3.7.3 (for implementation of experiments), TensorFlow 2.5.0 (for deep neural network), TensorFlow Probability 0.18.0 (for Flipout).



**Figure A1.** Architecture of the Bayesian neural network used in the experiment.

The BNN used in the experiment uses the evaluation factor extracted from the patent document as an input variable. Next, the input layer containing the factors passes the value of the patent to the output layer via the hidden layer. The weight of each layer is determined by the statistical distribution.

## References

1. Kaufmann, A.; Tödting, F. Science–Industry Interaction in the Process of Innovation: The Importance of Boundary-Crossing between Systems. *Res. Policy* **2001**, *30*, 791–804. [\[CrossRef\]](#)
2. Kim, J.; Lee, C.Y.; Cho, Y. Technological Diversification, Core-Technology Competence, and Firm Growth. *Res. Policy* **2016**, *45*, 113–124. [\[CrossRef\]](#)
3. Chen, Y.F.; Wu, T.C. An Empirical Analysis of Core Competence for High-Tech Firms and Traditional Manufacturers. *J. Manag. Dev.* **2007**, *26*, 159–168. [\[CrossRef\]](#)
4. Lin, B.-W.; Chen, C.-J.; Wu, H.-L. Patent Portfolio Diversity, Technology Strategy, and Firm Value. *IEEE Trans. Eng. Manag.* **2006**, *53*, 17–26. [\[CrossRef\]](#)
5. Appio, F.P.; De Luca, L.M.; Morgan, R.; Martini, A. Patent Portfolio Diversity and Firm Profitability: A Question of Specialization or Diversification? *J. Bus. Res.* **2019**, *101*, 255–267. [\[CrossRef\]](#)
6. Ha, S.H.; Liu, W.; Cho, H.; Kim, S.H. Technological Advances in the Fuel Cell Vehicle: Patent Portfolio Management. *Technol. Forecast. Soc. Chang.* **2015**, *100*, 277–289. [\[CrossRef\]](#)
7. Shi, X.; Cai, L.; Song, H. Discovering Potential Technology Opportunities for Fuel Cell Vehicle Firms: A Multi-Level Patent Portfolio-Based Approach. *Sustainability* **2019**, *11*, 6381. [\[CrossRef\]](#)
8. Chang, S. Bin Using Patent Analysis to Establish Technological Position: Two Different Strategic Approaches. *Technol. Forecast. Soc. Chang.* **2012**, *79*, 3–15. [\[CrossRef\]](#)
9. Kim, G.; Bae, J. A Novel Approach to Forecast Promising Technology through Patent Analysis. *Technol. Forecast. Soc. Chang.* **2017**, *117*, 228–237. [\[CrossRef\]](#)

10. Aristodemou, L.; Tietze, F. Citations as a Measure of Technological Impact: A Review of Forward Citation-Based Measures. *World Pat. Inf.* **2018**, *53*, 39–44. [[CrossRef](#)]
11. Falk, N.; Train, K. Patent Valuation with Forecasts of Forward Citations. *J. Bus. Valuat. Econ. Loss Anal.* **2017**, *12*, 101–121. [[CrossRef](#)]
12. Lanjouw, J.O.; Schankerman, M. Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators. *Econ. J.* **2004**, *114*, 441–465. [[CrossRef](#)]
13. Han, E.J.; Sohn, S.Y. Patent Valuation Based on Text Mining and Survival Analysis. *J. Technol. Transf.* **2015**, *40*, 821–839. [[CrossRef](#)]
14. Moore, K.A. Worthless Patents. *Berkeley Technol. Law J.* **2005**, *20*, 1521. [[CrossRef](#)]
15. Liu, K.; Arthurs, J.; Cullen, J.; Alexander, R. Internal Sequential Innovations: How Does Interrelatedness Affect Patent Renewal? *Res. Policy* **2008**, *37*, 946–953. [[CrossRef](#)]
16. Bozeman, B. Technology Transfer and Public Policy: A Review of Research and Theory. *Res. Policy* **2000**, *29*, 627–655. [[CrossRef](#)]
17. Cunningham, J.A.; Menter, M.; Young, C. A Review of Qualitative Case Methods Trends and Themes Used in Technology Transfer Research. *J. Technol. Transf.* **2016**, *42*, 923–956. [[CrossRef](#)]
18. Yoon, B.; Park, I.; Yun, D.; Park, G. Exploring Promising Vacant Technology Areas in a Technology-Oriented Company Based on Bibliometric Analysis and Visualisation. *Technol. Anal. Strateg. Manag.* **2019**, *31*, 388–405. [[CrossRef](#)]
19. Rotolo, D.; Hicks, D.; Martin, B.R. What Is an Emerging Technology? *Res. Policy* **2015**, *44*, 1827–1843. [[CrossRef](#)]
20. Noh, H.; Seo, J.H.; Sun Yoo, H.; Lee, S. How to Improve a Technology Evaluation Model: A Data-Driven Approach. *Technovation* **2018**, *72–73*, 1–12. [[CrossRef](#)]
21. Lee, J.; Kang, J.; Park, S.; Jang, D.; Lee, J. A Multi-Class Classification Model for Technology Evaluation. *Sustainability* **2020**, *12*, 6153. [[CrossRef](#)]
22. Jun, S. Text Data Analysis Using Generalized Linear Mixed Model and Bayesian Visualization. *Axioms* **2022**, *11*, 674. [[CrossRef](#)]
23. Uhm, D.; Jun, S. Zero-Inflated Patent Data Analysis Using Generating Synthetic Samples. *Futur. Internet* **2022**, *14*, 211. [[CrossRef](#)]
24. Lee, J.; Park, S. A Study on the Calibrated Confidence of Text Classification Using a Variational Bayes. *Appl. Sci.* **2022**, *12*, 9007. [[CrossRef](#)]
25. Choi, J.; Lee, C.; Yoon, J. Exploring a Technology Ecology for Technology Opportunity Discovery: A Link Prediction Approach Using Heterogeneous Knowledge Graphs. *Technol. Forecast. Soc. Chang.* **2023**, *186*, 122161. [[CrossRef](#)]
26. Woudenberg, F. An Evaluation of Delphi. *Technol. Forecast. Soc. Chang.* **1991**, *40*, 131–150. [[CrossRef](#)]
27. Kharat, M.G.; Raut, R.D.; Kamble, S.S.; Kamble, S.J. The Application of Delphi and AHP Method in Environmentally Conscious Solid Waste Treatment and Disposal Technology Selection. *Manag. Environ. Qual.* **2016**, *27*, 427–440. [[CrossRef](#)]
28. Galbraith, C.S.; Ehrlich, S.B.; DeNoble, A.F. Predicting Technology Success: Identifying Key Predictors and Assessing Expert Evaluation for Advanced Technologies. *J. Technol. Transf.* **2006**, *31*, 673–684. [[CrossRef](#)]
29. Akoka, J.; Comyn-Wattiau, I. A Method for Emerging Technology Evaluation. Application to Blockchain and Smart Data Discovery. In *Conceptual Modeling Perspectives*; Springer: Cham, Switzerland, 2017; pp. 247–258.
30. Sá, E.; Carvalho, A.; Silva, J.; Rezazadeh, A. A Delphi Study of Business Models for Cycling Urban Mobility Platforms. *Res. Transp. Bus. Manag.* **2022**, *45*, 100907. [[CrossRef](#)]
31. Woo Kim, H.; Kim, J.; Lee, J.; Park, S.; Jang, D. A Novel Methodology for Extracting Core Technology and Patents by IP Mining. *J. Korean Inst. Intell. Syst.* **2015**, *25*, 392–397. [[CrossRef](#)]
32. Choi, J.; Jang, D.; Jun, S.; Park, S. A Predictive Model of Technology Transfer Using Patent Analysis. *Sustainability* **2015**, *7*, 16175–16195. [[CrossRef](#)]
33. Kumari, R.; Jeong, J.Y.; Lee, B.-H.; Choi, K.-N.; Choi, K. Topic Modelling and Social Network Analysis of Publications and Patents in Humanoid Robot Technology. *J. Inf. Sci.* **2021**, *47*, 658–676. [[CrossRef](#)]
34. Lai, K.-K.; Chen, Y.-L.; Kumar, V.; Daim, T.; Verma, P.; Kao, F.-C.; Liu, R. Mapping Technological Trajectories and Exploring Knowledge Sources: A Case Study of E-Payment Technologies. *Technol. Forecast. Soc. Chang.* **2023**, *186*, 122173. [[CrossRef](#)]
35. Yang, D.S.-C.G. Some Methods Determining Reasonable Royalty Rates for Patent Valuation—An Infringement Damages Model. *J. Korea Technol. Innov. Soc.* **2012**, *15*, 700–721.
36. Trappey, A.J.C.; Trappey, C.V.; Wu, C.-Y.; Lin, C.-W. A Patent Quality Analysis for Innovative Technology and Product Development. *Adv. Eng. Inform.* **2012**, *26*, 26–34. [[CrossRef](#)]
37. Ko, N.; Jeong, B.; Seo, W.; Yoon, J. A Transferability Evaluation Model for Intellectual Property. *Comput. Ind. Eng.* **2019**, *131*, 344–355. [[CrossRef](#)]
38. Trappey, A.J.C.; Trappey, C.V.; Govindarajan, U.H.; Sun, J.J.H. Patent Value Analysis Using Deep Learning Models—The Case of IoT Technology Mining for the Manufacturing Industry. *IEEE Trans. Eng. Manag.* **2021**, *68*, 1334–1346. [[CrossRef](#)]
39. Chung, P.; Sohn, S.Y. Early Detection of Valuable Patents Using a Deep Learning Model: Case of Semiconductor Industry. *Technol. Forecast. Soc. Chang.* **2020**, *158*, 120146. [[CrossRef](#)]
40. Lee, C.-W.; Tao, F.; Ma, Y.-Y.; Lin, H.-L. Development of Patent Technology Prediction Model Based on Machine Learning. *Axioms* **2022**, *11*, 253. [[CrossRef](#)]
41. Huang, Z.; Li, J.; Yue, H. Study on Comprehensive Evaluation Based on AHP-MADM Model for Patent Value of Balanced Vehicle. *Axioms* **2022**, *11*, 481. [[CrossRef](#)]
42. Lecun, Y. A Theoretical Framework for Back-Propagation. In *Proceedings of the 1988 Connectionist Models Summer School*; Morgan Kaufmann: San Mateo, CA, USA, 1988; pp. 21–28.

43. Nair, V.; Hinton, G. Rectified Linear Units Improve Restricted Boltzmann Machines. In Proceedings of the ICML'10: 27th International Conference on Machine Learning, Haifa, Israel, 21–24 June 2010.
44. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
45. Gal, Y.; Ghahramani, Z. A Theoretically Grounded Application of Dropout in Recurrent Neural Networks. *Adv. Neural Inf. Process. Syst.* **2016**, *29*, 1–9.
46. Blundell, C.; Cornebise, J.; Kavukcuoglu, K.; Com, W. Weight Uncertainty in Neural Networks. In Proceedings of the ICML'15: 32nd International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 1613–1622.
47. Liu, Q.; Wang, D. Stein Variational Gradient Descent: A General Purpose Bayesian Inference Algorithm. *Adv. Neural Inf. Process. Syst.* **2016**, *29*, 1–9.
48. Graves, A. Practical Variational Inference for Neural Networks. *Adv. Neural Inf. Process. Syst.* **2011**, *24*, 1–9.
49. Wan, L.; Zeiler, M.; Zhang, S.; Cun, Y.L.; Fergus, R. Regularization of Neural Networks Using DropConnect. *PMLR* **2013**, *28*, 1058–1066.
50. Kingma, D.P.; Welling, M. Auto-Encoding Variational Bayes. *arXiv* **2013**. [[CrossRef](#)]
51. Kingma, D.P.; Salimans, T.; Welling, M. Variational Dropout and the Local Reparameterization Trick. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 1–9.
52. Wen, Y.; Vicol, P.; Ba, J.; Tran, D.; Grosse, R. Flipout: Efficient Pseudo-Independent Weight Perturbations on Mini-Batches. *arXiv* **2018**. [[CrossRef](#)]
53. Esteva, A.; Robicquet, A.; Ramsundar, B.; Kuleshov, V.; DePristo, M.; Chou, K.; Cui, C.; Corrado, G.; Thrun, S.; Dean, J. A Guide to Deep Learning in Healthcare. *Nat. Med.* **2019**, *25*, 24–29. [[CrossRef](#)]
54. Ding, X.; Nassehi, D.; Larson, E.C. Measuring Oxygen Saturation With Smartphone Cameras Using Convolutional Neural Networks. *IEEE J. Biomed. Health Inform.* **2019**, *23*, 2603–2610. [[CrossRef](#)] [[PubMed](#)]
55. Li, Q.; Ren, Z.; Cao, K.; Li, M.M.; Wang, K.; Zhou, Y. CancerVar: An Artificial Intelligence–Empowered Platform for Clinical Interpretation of Somatic Mutations in Cancer. *Sci. Adv.* **2022**, *8*, 1624. [[CrossRef](#)] [[PubMed](#)]
56. Shapiro, S.S.; Wilk, M.B. Biometrika Trust An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* **1965**, *52*, 591–611. [[CrossRef](#)]
57. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. *arXiv* **2017**. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.