Bayesian Graph Convolutional Neural Network based Patent Valuation Model

Weidong Liu
College of Computer Science,
Inner Mongolia University
Inner Mongolia Key laboratory of
Social Computing and Data Processing
Hohhot, China
cslwd@imu.edu.cn

Xin Liu
College of Computer Science,
Inner Mongolia University
Hohhot, China
imu.liuxx@gmail.com

Wenbo Qiao
College of Computer Science,
Inner Mongolia University
Hohhot, China
imu.qiaowb@gmail.com

***Abstract*—With the intense competition of global intellectual property, the increasing patents promote the potentiality of patent transactions. Patent valuation is the premise of the patent transaction. Automatic patent valuation faces some challenging issues from valuation feature to valuation model. To solve the above issues, we propose a Bayesian graph convolutional neural network based patent valuation model. In the model, the valuation objects are defined, from which to some valuation features are extracted. Valuation scenario is the constructed, on which Bayesian graph convolutional neural network is used to generate patent value. We evaluate our model by comparing the state-of-the-art model on patent data sets. The results show that our model outperforms other models in the evaluation measurements.**

***Index Terms*—patent valuation, Bayesian neural network**

I. Introduction

With the innovation-driven economic development, technological innovation and intellectual property (IP) are received unprecedented attention. As an important form of IP, patents play important roles in technology, law and economy. In economy, patents can be traded as special commodities. The patent value are being explored, which promotes the intellectual property economy. However, patent valuation research is still in the initial stage.

The current patent valuation methods are summarized into two categories:

1) Methods on valuation features [1]–[8]. Most methods extract some features from patent rather than other data for patent valuation.
2) Methods on valuation methods [9]–[20]. Most used models include regression based models, deep learning based models or probabilistic graph based models.

However, most methods measure patent value without the considerations of the valuation scenario. In fact, The patent data and its associated data often have many complex relations, which can be regards as the valuation scenarios. The patent value is relative with its valuation scenarios. The patent value varies with different valuation scenarios and thus is relative with its valuation scenarios.

Herein, we propose a Bayesian graph convolutional neural network based patent valuation model.

In the model, the patent valuation is modeled by a probabilistic generative process on Bayesian graph convolutional neural network, resulting in posterior distributions of patent value.

II. Preliminaries and Problem Statement

A. Primary Knowledge

Valuation objects denote the some objects to be valued.

Definition 1: Valuation Objects(VO)

$$VO = \{o_n\}_{n=1}^{|VO|} \tag{1}$$

where o denotes a object to be valued; t^o denotes the type of o; $T = \{t_m\}_{m=1}^{|T|}$ denotes different object types and $t^o \in T$.

Definition 2: Valuation Scenario(VS)

$$VS = \{VO, R\} \tag{2}$$

$R = \{r^{(o,o^-)} | o \in VO \ and \ o < o^-\}$ denotes set of weights between valuation objects;

For the pair (o, o^-), $r^{(o,o^-)}$ denotes a weight 1 if there is a association relation between o and o^-, 0 otherwise;

The valuation scenario is generated by the following steps [21]:

1) draw $m^{(t,t^-)} \sim Gam(\alpha, \beta)$ for $M = \{m^{(t,t^-)} | t \in T, t < t^-\}$
2) draw $\phi^z_{k1} \sim Gam(\alpha, \beta)$ for $\{\phi^z_i | z \in Z\}_{k1=1}^{K1}$
3) draw $\phi^o_{k2} \sim Gam(\alpha^o, \beta^o)$ for $\{\phi^o_{k2} | o \in VO\}_{k2=1}^{K2}$
4) For each object pairs $(o, o^-) \in \{(o, o^-) | o \in VO, o < o^-\}$,
   a) draw $z^{(o,o^-)} \sim \pi^o$ where $\pi^o_i = \frac{\phi^o_i}{\sum \phi^o_i}$ denotes a K dimension probability distribution of o participating in the K semantic communities.
   b) draw $z^{(o^-,o)} \sim \pi^{o^-}$ where $\pi^{o^-}_i = \frac{\phi^{o^-}_i}{\sum \phi^{o^-}_i}$
   c) if $z^{(o,o^-)} = z^{(o^-,o)}$, draw $r^{(o,o^-)} \sim Bern(\frac{m^{(t^o,t^{o^-})}\phi^z_1}{m^{(t^o,t^{o^-})}\phi^z_1 + \phi^z_2})$
   Otherwise $r^{(o,o^-)} \sim Bern(\delta)$

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

The probability distribution of $VS = \{VO, R\}$ is computed by,

$$p(VS) = \sum_{M} \sum_{Phi} p(VS|\Phi, M)p(M)p(\Phi) \tag{3}$$

B. Problem statement

Given the features $X = \{x^o | o \in VO - VO^*\}$, the distribution of their values $Y = \{y^o | o \in VO - VO^*\}$ is defined by,

$$\begin{aligned} p(Y|X, VS^*, VO^*) = \int & p(Y|X, W, VS, \Gamma_Y)p(W, \Gamma|VO^*, VS) \\ & P(VS|\Phi, M)P(\Phi, M|VS^*) \\ & dVSd\Phi dWd\Gamma dM \end{aligned} \tag{4}$$

and

$$p(Y|X, W, VS, \Gamma_Y) = \prod_{o \in VO} \prod_{i=1}^{K} p(y_i^o | x^o, W^{t^o}, VS, \gamma_Y^{t^o}) \tag{5}$$

and

$$p(\Phi, M|VS^*) = p(\Phi|VS^*)p(M|VS^*) \tag{6}$$

and

$$p(W, \Gamma|VO^*, VS) = \frac{p(Y^*|X^*, W, VS, \Gamma_Y)p(W|\Gamma_W)p(\Gamma)}{p(Y^*|X^*)} \tag{7}$$

and

$$p(W|\Gamma_W) = \prod_{t \in T} p(W^t | \gamma_W^t) \tag{8}$$

and

$$p(\Gamma) = \prod_{t \in T} p(\gamma^t) \tag{9}$$

where $Y = \{y^o | o \in VO - VO^*\}$; $X = \{x^o | o \in VO - VO^*\}$; $R = \{r^{(o,o^-)}\}$ denotes a set of relation among VO; $W = \{W^t | t \in T\}$ denotes a set of weights for T; $\Gamma = \{\gamma^t | t \in T\}$ denotes a set of precise parameters; VS denotes a valuation scenario to be generated; VS^* denotes a generated valuation scenario.

We list the notations of the above definitions in TableI, which are thoroughly used in this paper.

TABLE I
SYMBOLS AND THEIR MEANINGS

symbols	Description
$VO = \{o_n\}_{n=1}^{\lvert VO \rvert}$	valuation objects
$VO^* = \{o_n^*\}_{n=1}^{\lvert VO^* \rvert}$	marked valuation objects
$VO - VO^*$	unmarked valuation objects
x^o	features of o
y^o	values of o
t^o	type of o
$T = \{t_m\}_{m=1}^{\lvert T \rvert}$	object types
$Y^* = \{y^o \lvert o \in VO^*\}$	marked values of VO^*
$X^* = \{x^o \lvert o \in VO^*\}$	features of VO^*
$Y = \{y^o \lvert o \in VO - VO^*\}$	values of $VO - VO^*$
$X = \{x^o \lvert o \in VO - VO^*\}$	features of $VO - VO^*$
$r^{(o,o^-)}$	weight of (o, o^-)
$R = \{r^{o,o^-} \lvert o \in VO, o < o^-\}$	relations among VO
$VS = \{VO, R\}$	valuation scenario
$VS^* = \{VO, R\}$	observable valuation scenario
$L = \{l_i\}_{i=1}^{\lvert L \rvert}$	the layers of weight matrix
$W_l = \{w_{i,j,l} \lvert l \in L\}$	weight matrix of the l^{th} layer
$W = \{W_l^t \lvert t \in T, l \in L\}$	L weight matrix for types T
γ_W^t	precise parameter of W
γ_Y^t	precise parameter of Y
$\gamma^t = \{\gamma_W^t, \gamma_Y^t \lvert t \in T\}$	precise parameters for t
$\Gamma = \{\gamma^t \lvert t \in T\}$	a set of precise parameters for T
$M = \{m^{(t,t^-)} \lvert t \in T \ and \ t < t^-\}$	the set of intensity among T
$\Phi = \{\phi_i^o \lvert o \in VO \ and \ 1 \leq i \leq K\}$	community-structural parameters

III. Proposed Model and Inference

A. Proposed Model

For the valuation objects with the concrete type t, a Bayesian graph convolutional neural network based patent valuation model is proposed by,

Definition 3:Valuation Model, (VM)

$$VM = (Y^t, W^t | X^t, VO^*, VS^*, \Gamma^*, \Phi^*, M^*) \tag{10}$$

and

$$M^*, \Phi^* = \arg\max_{M,\Phi} P(M, \Phi|VS^*) \tag{11}$$

and

$$\begin{aligned} \Gamma_{W_l^t}^* &= \{\gamma_{W_l^t}^*(x, y) | l \in L, x \in h_l, y \in h_l\} \\ \gamma_{W_l^t}^*(x, y) &= |N_l|^{-1}\sigma(x)\sigma(y) \end{aligned} \tag{12}$$

and

$$\begin{aligned} \Gamma_{Y^t}^* &= \{\gamma_{Y^t}^*(y) | y \in Y\} \\ \gamma_{Y^t}^*(y) &= \tau^{-1}\mathrm{I} \end{aligned} \tag{13}$$

Fig1 shows the graph representation of VM, in which some symbols and their meanings are listed in Table.I.

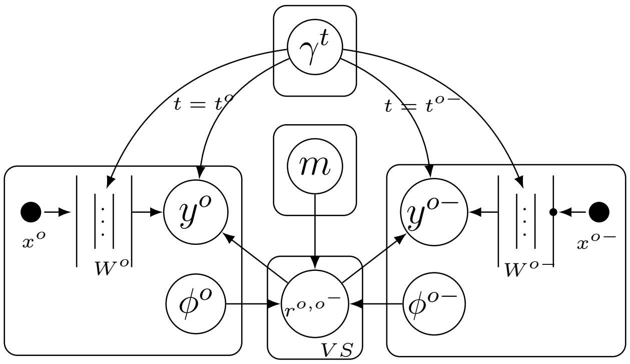

Fig. 1. Association Patent Valuation Model

Given VM, the distribution of multi-dimensional patent values is obtained by Eq.14.

$$\begin{aligned} p(Y^t|X^t, VS^*, VO^*, \Gamma^{t*}, \Phi^*, M^*) = \int & p(Y^t|X^t, W, \Gamma^*) \\ & p(W^t|VO^*, VS) \\ & P(VS|\Phi^*, M^*) \\ & dVSdW^t \end{aligned} \tag{14}$$

Given $X^t, VO^*, VS^*, \Gamma^*, \Phi^*, M^*$, the patent value is formed by,

1) draw $VS \sim p(VS|\Phi^*, M^*)$
2) draw $W_l^t \sim p(W_l^t|VO^*, VS)$ for $W = \{W_l^t|t \in T, l \in L\}$
3) draw $Y^t \sim N(f(X^t; W^t; VS), {\Gamma_Y^t}^{-1})$ for $Y = \{Y^t\}$.

where

$$\begin{aligned} f(X^t; W^t; VS) &= h_l \\ h_1 &= \sigma(A^{VS} X W_1^t) \\ h_l &= \sigma(A^{VS} h_l W_l^t) \end{aligned} \tag{15}$$

B. Inference

Monte Carlo approximation is used in Eq.14 [20]. The patent value in is computed by,

$$p(Y|X, VS^*, VO^*) = \frac{1}{|\widetilde{VS}|} \frac{1}{|\widetilde{W}|} \sum_{vs_i \in \widetilde{VS}} \sum_{w_{i,j} \in \widetilde{W}} p(Y|X, w_{i,j}, vs_i, \Gamma^*) \tag{16}$$

where $\widetilde{VS} = \{vs_i\}_{k=i}^{|\widetilde{VS}|}; \widetilde{W} = \{w_j\}_{j=1}^{|\widetilde{W}|}$.

$$vs_i \sim P(vs_i|\Phi_j^*, M_i^*) \tag{17}$$

and

$$w_{i,j} \sim P(w_{i,j}|vs_i, VO^*) \tag{18}$$

and

Algorithm 1: Valuation Model

Input: VO, VS^*, VO^*
Output: $p(Y|X, VS^*, VO^*)$
Initialization:;
Obtaining Φ^* and M^*;
for $i = 1 : |\widetilde{VS}|$ **do**
  sample $vs_i \sim P(vs_i|\Phi^*, M^*)$;
  for $j = 1 : |\widetilde{W}|$ **do**
    sample $w_{i,j} \sim P(w_{i,j}|vs_i, \Gamma^*, VO^*)$;
  end
end
approximate $p(Y|X, VS^*, VO^*)$ by Eq.16

IV. Experiments

To validate efficiency of our valuation model, we design experiments as follows.

A. Experimental Datasets

The company annual reports and patents used in our experiments are shown in table II.

The dataset 1 includes 1187 patents of internal patent classification(IPC) A61 and 232 Chinese listed companies annual reports of industry classification(IC) 27.

The dataset 2 includes 2742 patents of internal patent classification(IPC) H01 and 244 Chinese listed companies annual reports of industry classification(IC) 38.

TABLE II
The description of experimental data

Data Set	Dataset 1	Dataset 2
#IPC	A61	H01
total number of patent	1187	2742
#IC	27	38
total number of report	232	244

The datasets are collected by the following steps,

1) Download 1187 patents under the IPC A61 and 2742 patents under the IPC H01 under IPC H01 from the State Intellectual Property Office of China(SIPO)[1].
2) Download industry classification(IC) standard from China Securities Regulatory Commission.
3) Map IPC A61 and IPC H01 to their most similar code of industry classification(IC) 27 and 38 respectively.
4) Download 232 reports under the IC 27 and 244 reports under IC 38 from CNINF[2].

B. Baseline Models

We compare our model with the Bayesian neural network(BNN) [22] based model in our experiments.

[1] http://www.cnipa.gov.cn/
[2] http://www.cninfo.com.cn/new/index

C. Evaluation Measurements

Given the valuation objects $VO = \{o_n\}_{n=1}^{|VO|}$, Mean Absolute Error(MAE) and Mean Relative Error(MRE) are used to compare the results of the model with the benchmark data by,

$$MAE = \frac{1}{|VO|} \sum_{o \in VO} |y^o - \bar{y^o}| \tag{19}$$

$$MRE = \frac{1}{|VO|} \sum_{o \in VO} |(y^o - \bar{y^o})/\bar{y^o}| \tag{20}$$

where $Y = \{y^o | o \in VO\}$ denotes the patent values obtained by the automatic model; $\bar{Y} = \{\bar{y}^o | o \in VO\}$ denotes the benchmark of patent value or patent reference values.

D. Experimental Setups

We conduct experiment on dataset 1 and dataset 2 respectively.

1) Construct a valuation scenario VS by,
   a) Embed textual parts of $o \in VO$ by doc2vec(e.g. the abstract in patent or the main business in annual reports);
   b) Compute $r^{(o,o^-)}$ for each pair of patent and annual report;
   c) Keep $R = \{r^{(o,o^-)} | r^{(o,o^-)} > \theta\}$;
   d) Link o and o^- to $r^{(o,o^-)}$;
   e) Form $VS = \{VO, R\}$.
2) Extract a feature vector $x^o = (x_i^o)_{i=1}^M$ for each object $o \in VO$. For patents and company annual report in table III.
   a) if the feature is included in the object type t^o, the feature values is computed respectively;
   b) if the feature is not included in the object type t^o, the feature value is set 0.
3) Compute the posterior distribution of patent value in Eq.16;
4) Compare the results of our model with that of baseline models in the section IV-B;
5) Evaluate the results on the evaluation measurement in the section IV-C.

TABLE III
FEATURES EXTRACTED FROM VALUATION OBJECTS

Objects Types	Features
patent	number of claims number of patent family number of classification time since authorization
company annual report	earning rate of net assets growth rate of operating income growth rate of net profit rate of gross profit asset liability ratio account receivable turnover

E. Experimental Results

We compare our model on the dataset 1 and dataset 2 with the baseline model in section IV-B under the measurements in section IV-C. Since there are no standard patent reference values, the patent reference values are replaced by the number of forward citation.

We compare the mean absolute error(MAE) and the mean relative error(MRE) of our model with that of the basic models on two datasets. Table IV, Fig.2-Fig.5 show the comparative results.

Fig.2 and Fig.3 show that our model has lower value in the mean absolute error(MAE) and on the mean relative error(MRE) for the dataset 1 with A61. It indicates that our model outperforms the baseline model on the precision including MAE and MRE for the dataset 1 with A61.

TABLE IV
MAE AND MRE COMPARISON BETWEEN OUR MODEL AND THE BASELINE MODEL

domain	measurement	BNN	our model
Dataset 1(A06)	MAE	60.094	110.677
	MRE	0.591	1.682
Dataset 2(H01)	MAE	35.000	88.616
	MRE	1.383	2.845

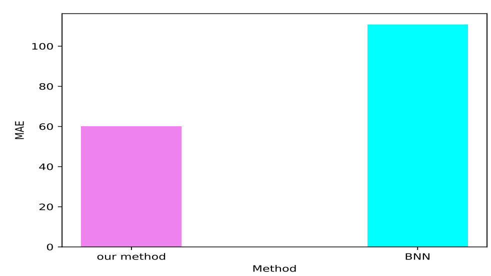

Fig. 2. MAE of BNN and our model on dataset1(A61)

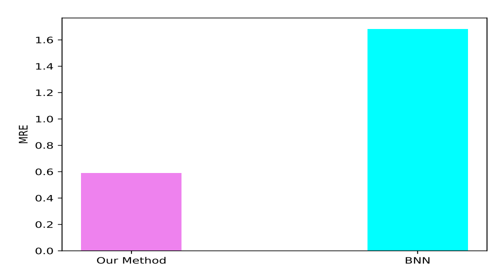

Fig. 3. MRE of BNN and our model on dataset1(A61)

Fig.4 and Fig.5 show that our model has lower value in the mean absolute error(MAE) and on the mean relative error(MRE) for the dataset 2 with H01. Our model outperforms the baseline model on MAE and MRE for the dataset 2 with H01.

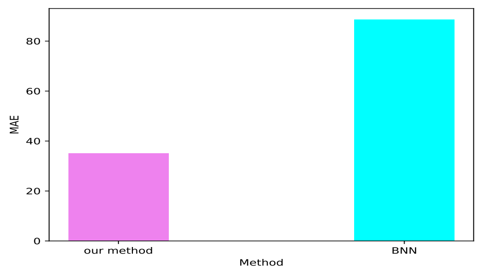

Fig. 4. MAE of BNN and our model on dataset2(H01)

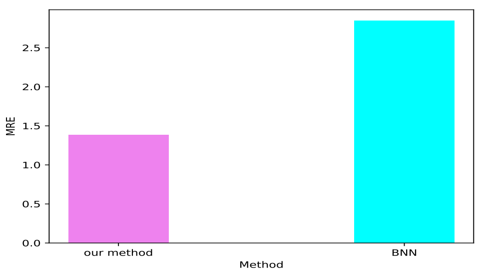

Fig. 5. MRE of BNN, Our model on dataset2(H01)

We compare the change of mean relative error(MRE) with the patent value in range 0-300 on dataset1(A61) and dataset2(H01).

The comparative results of dataset1(A61) are shown in tableV and Fig.6. Compared with that of the baseline model, our model has lower mean relative error on different patent values.

TABLE V
THE CHANGE OF MRE WITH DIFFERENT VALUES ON DATASET1(A61)

value range	BNN	our model
(0-50]	2.972	0.922
(50-100]	1.168	0.212
(100-150]	0.76	0.540
(150-200]	0.816	0.677
(200-250]	0.816	0.767
(250-300]	0.900	0.820

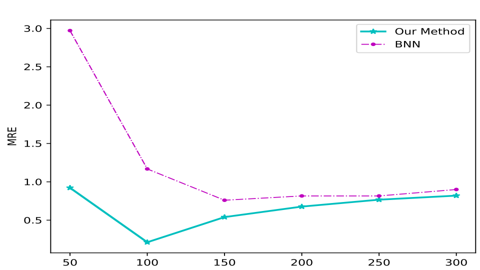

Fig. 6. MRE of BNN and our model change with value on dataset1(A61)

The comparative results of dataset1(H01) are shown in tableVI and Fig.7. The mean relative error obtained by our methods is lower than that of the baseline model on different patent values.

TABLE VI
THE CHANGE OF MRE WITH DIFFERENT VALUES ON DATASET2(H01)

value range	BNN	our model
(0-50]	4.46	2.290
(50-100]	1.250	0.280
(100-150]	0.950	0.580
(150-200]	0.825	0.700
(200-250]	0.848	0.793
(250-300]	0.760	0.820

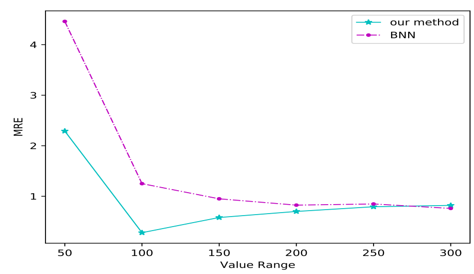

Fig. 7. MRE of BNN and Our model change with value on H01

The above results show that the addition of valuation scenario can improve the precision of the patent valuation, resulting lower mean absolute error and the mean relative error.

V. CONCLUSIONS AND FUTURE WORK

The paper proposes a patent valuation model to measure patent relative value in its valuation scenario. In our model, the patent value is a multi-dimensional vector, where each dimension exhibits a relative value for its its valuation scenario. a Bayesian graph convolutional neural network based model is used to discover the distributions of patent value which is a posterior distribution of the valuation scenario. To evaluate our model, our model is compared with state-of-the-art model on the patent datasets. The results show that our model outperforms other models in patent valuation.

ACKNOWLEDGMENT

The research work in this paper was supported by the National Science Foundation of China (grant no. 61801251) and Natural Science Foundation of Inner Mongolia (2018B-S06002).

REFERENCES

[1] T. Fischer and J. Leidinger, "Testing patent value indicators on directly observed patent valuetion: An empirical analysis of ocean tomo patent auctions," *Research Policy*, vol. 43, no. 3, pp. 519–529, 2014.

[2] X. Liu, J. Yan, S. Xiao, X. Wang, H. Zha, and S. M. Chu, "On predictive patent valuation: Forecasting patent citations and their types," in *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

[3] M. de Saint-Georges and B. v. P. de la Potterie, "A quality index for patent systems," *Research Policy*, vol. 42, no. 3, pp. 704–719, 2013.

[4] S. Wang, Z. Lei, and W.-C. Lee, "Exploring legal patent citations for patent valuation," in *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*. ACM, 2014, pp. 1379–1388.

[5] E. J. Han and S. Y. Sohn, "Patent valuation based on text mining and survival analysis," *The Journal of Technology Transfer*, vol. 40, no. 5, pp. 821–839, 2015.

[6] M. Grimaldi and L. Cricelli, "Indexes of patent value: a systematic literature review and classification," *Knowledge Management Research & Practice*, pp. 1–20, 2019.

[7] G. Thoma, "Composite value index of patent indicators: Factor analysis combining bibliographic and survey datasets," *World patent information*, vol. 38, pp. 19–26, 2014.

[8] S. Bass and L. Kurgan, "Discovery of factors influencing patent value based on machine learning in patents in the field of nanotechnology," *Scientometrics*, vol. 82, no. 2, pp. 217–241, 2009.

[9] X. Wang, H. Zhang, and Y. Liu, "Sentence vector model based on implicit word vector expression," *IEEE Access*, vol. 6, pp. 17 455–17 463, 2018.

[10] X. Wang, S. Yuan, H. Zhang, and Y. Liu, "Estimation of inter-sentiment correlations employing deep neural network models." *arXiv preprint arXiv:1811.09755*, 2018.

[11] J. Suzuki, "Structural modeling of the value of patent," *Research Policy*, vol. 40, no. 7, pp. 986–1000, 2011.

[12] S. Ercan and G. Kayakutlu, "Patent value analysis using support vector machines," *Soft computing*, vol. 18, no. 2, pp. 313–328, 2014.

[13] H. Lin, H. Wang, D. Du, H. Wu, B. Chang, and E. Chen, "Patent quality valuation with deep learning models," in *International Conference on Database Systems for Advanced Applications*. Springer, 2018, pp. 474–490.

[14] Y. Liu, P.-y. Hseuh, R. Lawrence, S. Meliksetian, C. Perlich, and A. Veen, "Latent graphical models for quantifying and predicting patent quality," in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2011, pp. 1145–1153.

[15] F. R. Denton and P. J. Heald, "Random walks, non-cooperative games, and the complex mathematics of patent pricing," *Rutgers L. Rev.*, vol. 55, p. 1175, 2002.

[16] W. J. Murphy, J. L. Orcutt, and P. C. Remus, *Patent Valuation: improving decision making through analysis*. John Wiley & Sons, 2012, vol. 571.

[17] J. R. Allison, M. A. Lemley, K. A. Moore, and R. D. Trunkey, "Valuable patents," *Geo. Lj*, vol. 92, p. 435, 2003.

[18] M. A. Lemley and C. Shapiro, "Probabilistic patents," *Journal of Economic Perspectives*, vol. 19, no. 2, pp. 75–98, 2005.

[19] F. Munari and R. Oriani, *The economic valuation of patents: methods and applications*. Edward Elgar Publishing, 2011.

[20] M. Coates, S. Pal, Y. Zhang, and D. Ustebay, "Bayesian graph convolutional neural networks for semi-supervised classification," *AAAI 2019 : Thirty-Third AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 5829–5836, 2019.

[21] W. Li, S. Ahn, and M. Welling, "Scalable mcmc for mixed membership stochastic blockmodels," in *Artificial Intelligence and Statistics*, 2016, pp. 723–731.

[22] Y. Gal and Z. Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning," in *international conference on machine learning*, 2016, pp. 1050–1059.