شبکه‌های عصبی مصنوعی

درس ۲۴

مورد مطالعاتی ۲:

تخمین احتمال

Case Study 2: Probability Estimation

کاظم فولادی قلعه

دانشکده مهندسی، پردیس فارابی

دانشگاه تهران

http://courses.fouladi.ir/nn

Probability Estimation Case Study: Molecular Dynamics

شبکه‌های عصبی مصنوعی

مورد مطالعاتی ۲ : تخمین احتمال

۱

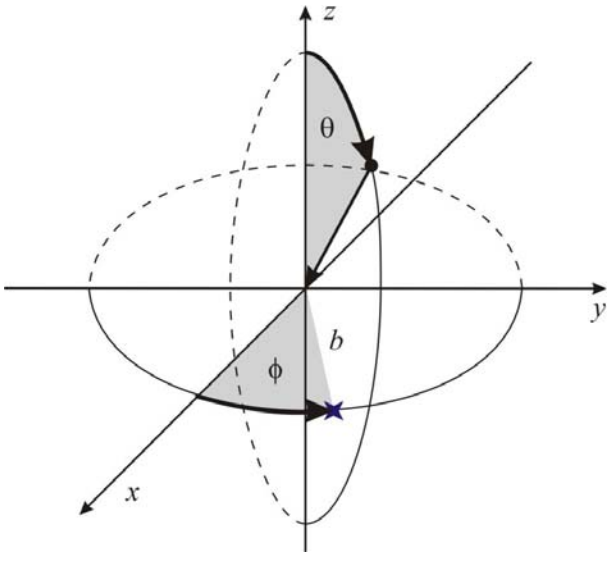
توصیف فرآیند CVD

Carbon Vapor Deposition (CVD)

- A carbon dimer is projected toward a diamond substrate.
- We will assume that the dimer can react with the substrate in one of three ways:
  - chemisorption
    (the atoms in the dimer become bound to the substrate),
  - scattering
    (the atoms bounce off the substrate),
  - desorption
    (the atoms become bound to the substrate for a period of time, but are then released).

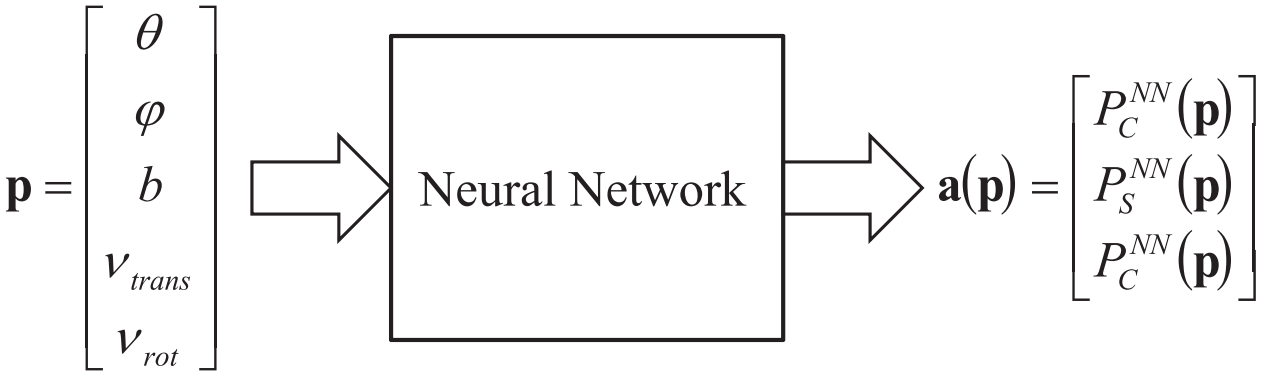
Variable Definitions

- Black circle represents the carbon dimer,
- Directed line represents the direction of the initial velocity vector.
- Blue star represents the location of the central carbon atom in the diamond substrate.
- Angle θ denotes the angle of incidence, i.e., the angle between the direction of the initial velocity vector of the carbon dimer and the perpendicular on the surface (the *z* direction).



- Impact parameter *b* is the distance between the location of the central atom and the point of intersection of the initial velocity vector and the diamond surface (the origin of the axes).
- Angle ϕ represents the angle between the *x* axis and the line from the origin to the central atom.

Neural Network



ν_{trans} - translational velocity of the C_2 dimer

ν_{rot} - rotational velocity of the C_2 dimer

$P_C^{NN}(\mathbf{p})$ - NN prediction of chemisorption probability

$P_S^{NN}(\mathbf{p})$ - NN prediction of scattering probability

$P_D^{NN}(\mathbf{p})$ - NN prediction of desorption probability

شبکه‌های عصبی مصنوعی

مورد مطالعاتی ۲ : تخمین احتمال

۲

گردآوری داده‌ها و پیش‌پردازش

Monte Carlo Data Collection

- Data for training the neural network are obtained by molecular dynamics (MD) simulations, where the motion of atoms and molecules in a material under a given force are simulated, using known laws of physics to calculate the forces on individual atoms
- We use a total of 324 atoms to model the CVD system. Out of these, 282 atoms of diamond substrate are used to model the crystalline face with 40 atoms of hydrogen on the top layer of the diamond surface, and 2 atoms in the C_2 dimer.
- The term Monte Carlo refers to the set of simulations that are obtained by setting a number of the variables to random values for each trajectory. We refer to the simulation of a single trajectory as an MD simulation, since the principles of molecular dynamics are used to perform the computations.

Training Data

The targets are obtained by estimating the probabilities of chemisorption, scattering or desorption from the Monte Carlo trials:

$$P_X^{MC}(\mathbf{p}) = \frac{N_X}{N_T}$$

where N_X is the number of MD trajectories that resulted in reaction X, and N_T is the total number of MD trajectories computed in the Monte Carlo trials. Since we do not know the true underlying reaction probabilities, we use the estimates obtained from the Monte Carlo trials as target outputs for the neural network. We can think of these estimates as noisy versions of the true probabilities.

2000 different input/target pairs were generated. Of these, 70% were randomly selected for training, 15% for validation, and 15% for testing. For each trajectory, the **p** were generated randomly, using physically-appropriate distributions for each variable. A total of 50 different trajectories were run to obtain each estimated probability. This means that 2000x50 trajectories were run to create the entire data set.

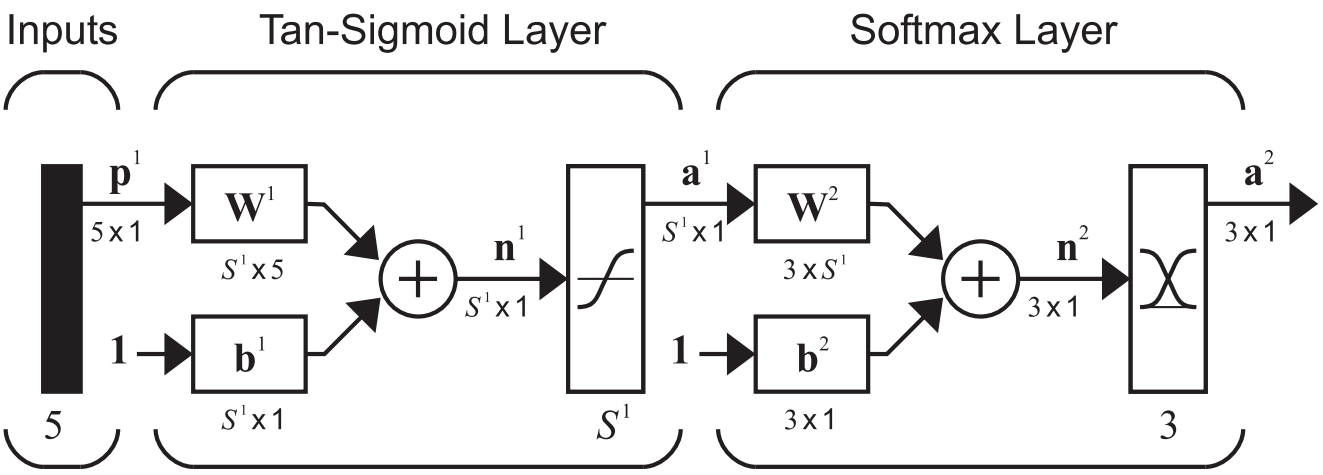
شبکه‌های عصبی مصنوعی

مورد مطالعاتی ۲ : تخمین احتمال

۳

انتخاب معماری

Network Architecture



$$\mathbf{a}^1 = \mathbf{tansig}(\mathbf{W}^1\mathbf{p} + \mathbf{b}^1) \qquad \mathbf{a}^2 = \mathbf{softmax}(\mathbf{W}^2\mathbf{a}^1 + \mathbf{b}^2)$$

$$\mathbf{t} = \begin{bmatrix} P_C^{MC}(\mathbf{p}) \\ P_S^{MC}(\mathbf{p}) \\ P_C^{MC}(\mathbf{p}) \end{bmatrix}$$

Softmax Transfer Function

$$a_i = f(n_i) = \frac{\exp(n_i)}{\sum_{j=1}^{S} \exp(n_j)}$$

$$\dot{\mathbf{F}}^m(\mathbf{n}^m) = \begin{bmatrix} a_1^m\left(\sum_{i=1}^{S^m} a_i^m - a_1^m\right) & -a_1^m a_2^m & \cdots & -a_1^m a_{S^m}^m \\ -a_2^m a_1^m & a_2^m\left(\sum_{i=1}^{S^m} a_i^m - a_2^m\right) & \cdots & -a_2^m a_{S^m}^m \\ \vdots & \vdots & \ddots & \vdots \\ -a_{S^m}^m a_1^m & -a_{S^m}^m a_2^m & \cdots & a_{S^m}^m\left(\sum_{i=1}^{S^m} a_i^m - a_{S^m}^m\right) \end{bmatrix}$$

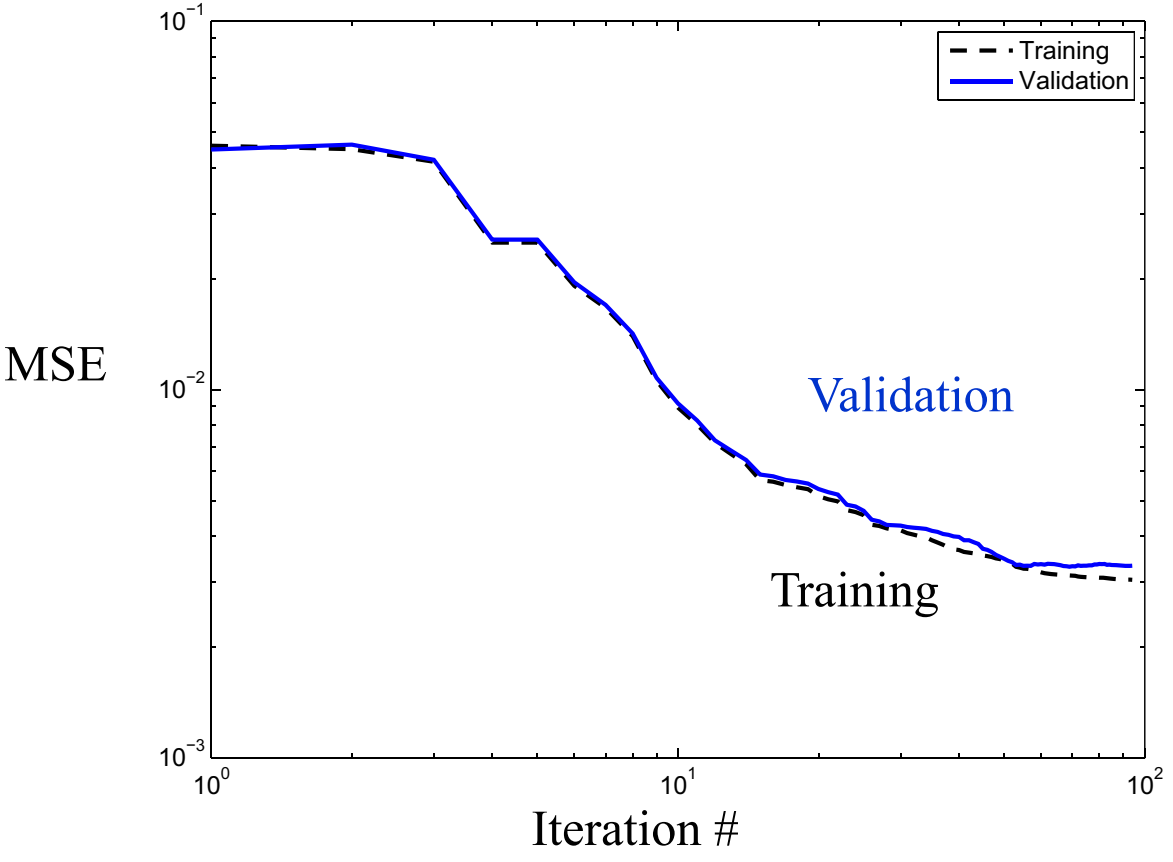
شبکه‌های عصبی مصنوعی

مورد مطالعاتی ۲ : تخمین احتمال

۴

آموزش شبکه

Training Performance (S^1=10)



Training and Validation Performance

$S^1 = 10$

	Training RMSE	Validation RMSE
$P_C(\mathbf{p})$	0.0496	0.0493
$P_S(\mathbf{p})$	0.0634	0.0659
$P_D(\mathbf{p})$	0.0586	0.0604

$S^1 = 2$

	Training RMSE	Validation RMSE
$P_C(\mathbf{p})$	0.0634	0.0627
$P_S(\mathbf{p})$	0.0669	0.0704
$P_D(\mathbf{p})$	0.0617	0.0618

$S^1 = 20$

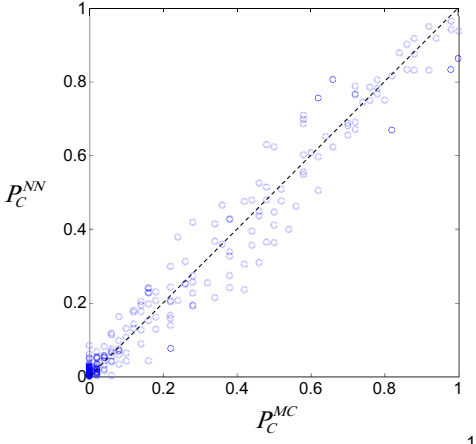
	Training RMSE	Validation RMSE
$P_C(\mathbf{p})$	0.0432	0.0444
$P_S(\mathbf{p})$	0.0603	0.0643
$P_D(\mathbf{p})$	0.0569	0.0595

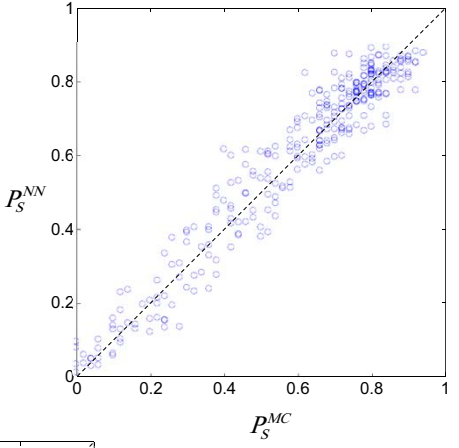
Multiple Initial Conditions

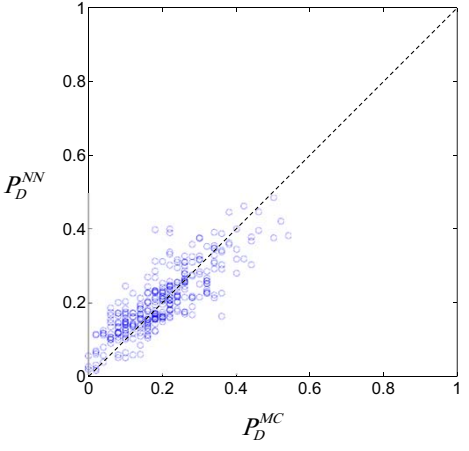
3.046e-003	2.953e-003	3.031e-003	3.105e-003	3.050e-003

- Final validation MSE for five different training runs.
- All of the errors are similar, so we have reached a global minimum at each run.
- If one error was significantly lower than the others, then we would use the weights that obtained the lowest error.

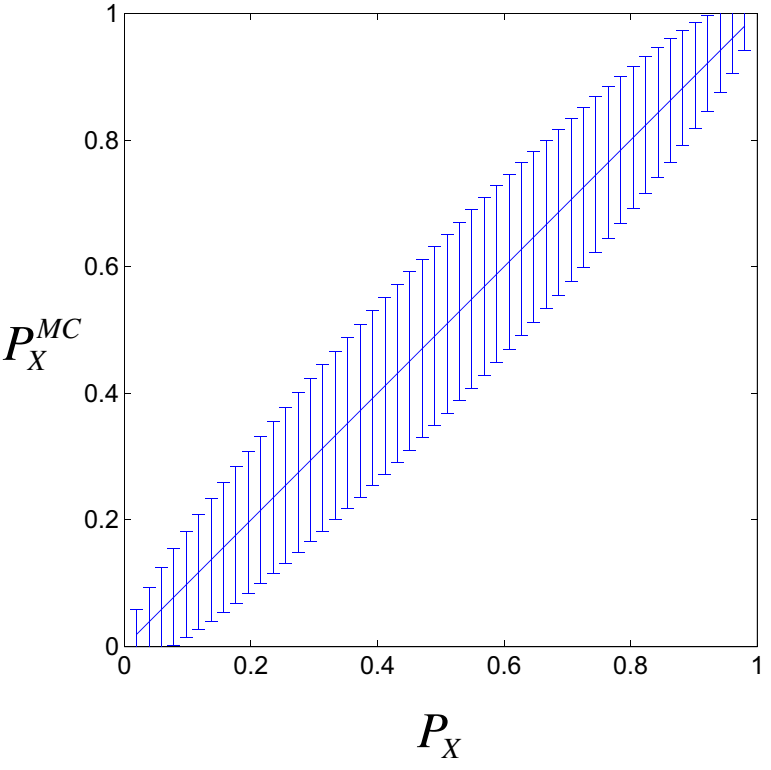
Output vs Target Scatter Plots







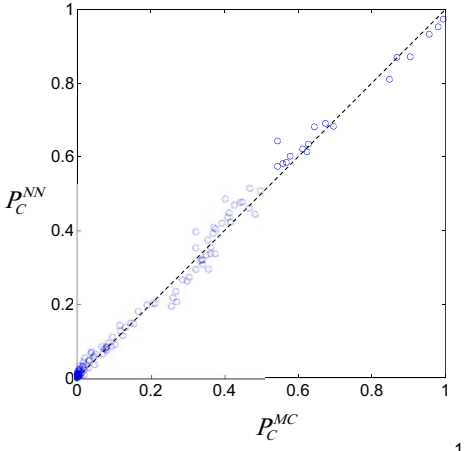
Statistical Error in P_X^{MC}

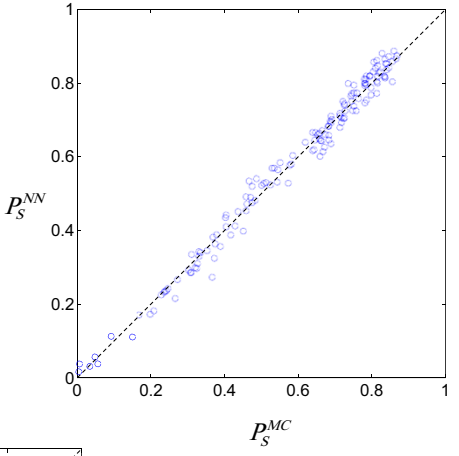


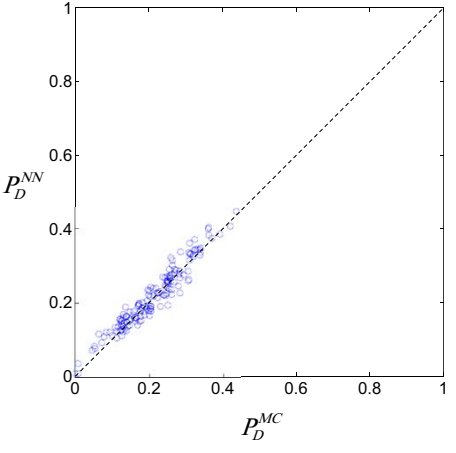
$$P_X - 2\Delta \le P_X^{MC} \le P_X + 2\Delta$$

$$\Delta = \sqrt{\frac{P_X(1-P_X)}{N_T}}$$

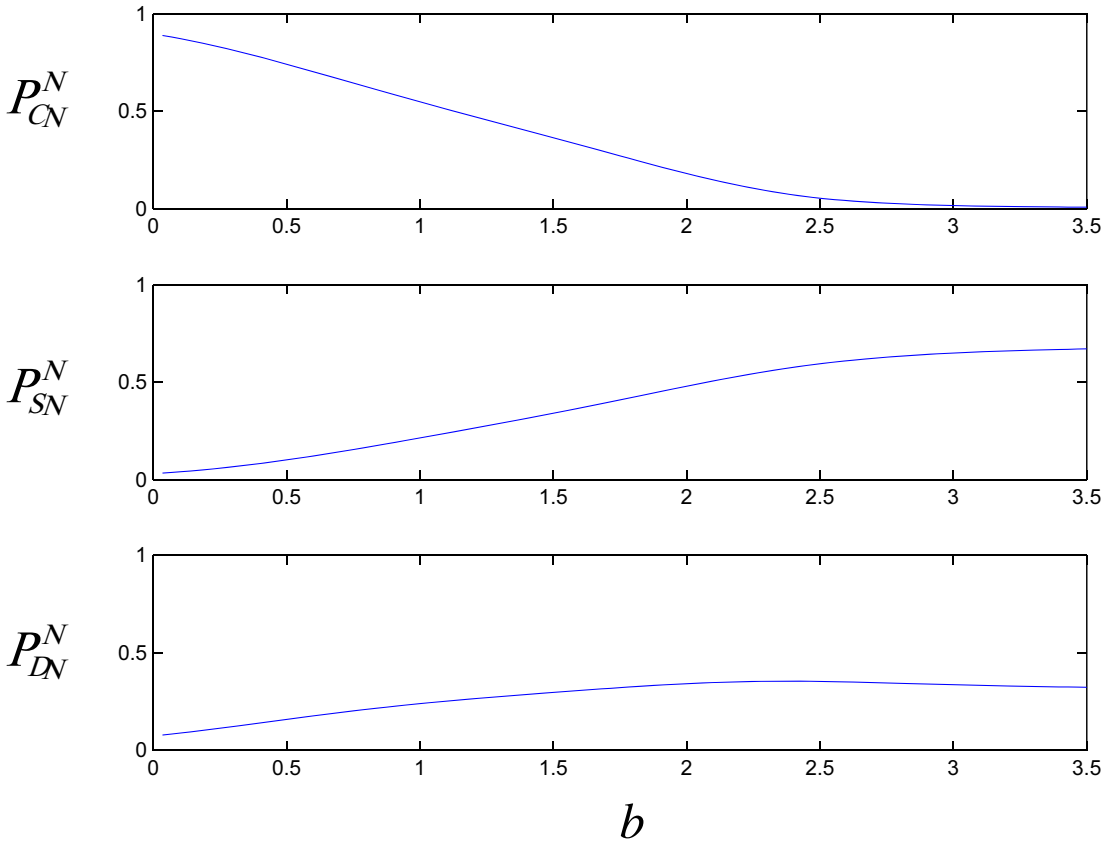
Outputs vs Targets for $N_T = 500$







Reaction Prob. vs Impact Parameter

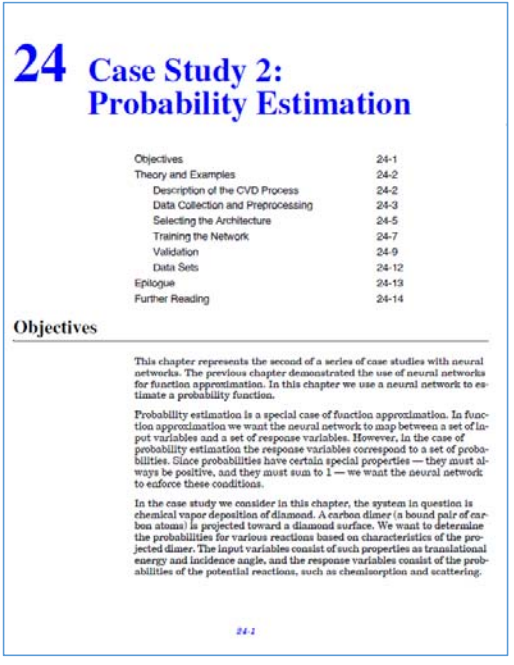


مورد مطالعاتی ۲ : تخمین احتمال

۵

منابع

منبع اصلی



24 Case Study 2: Probability Estimation

Objectives

This chapter represents the second of a series of case studies with neural networks. The previous chapter demonstrated the use of neural networks for function approximation. In this chapter we use a neural network to estimate a probability function.

Probability estimation is a special case of function approximation. In function approximation we want the neural network to map between a set of input variables and a set of response variables. However, in the case of probability estimation the response variables correspond to a set of probabilities. Since probabilities have certain special properties — they must always be positive, and they must sum to 1 — we want the neural network to enforce these conditions.

In the case study we consider in this chapter, the system in question is chemical vapor deposition of diamond. A carbon dimer (a bound pair of carbon atoms) is projected toward a diamond surface. We want to determine the probabilities for various reactions based on characteristics of the projected dimer. The input variables consist of such properties as translational energy and incidence angle, and the response variables consist of the probabilities of the potential reactions, such as chemisorption and scattering.

24-1

Martin T. Hagan, Howard B. Demuth, Mark H. Beale, Orlando De Jesus,
Neural Network Design,
2nd Edition, Martin Hagan, 2014.
Chapter 24

Online version can be downloaded from: http://hagan.okstate.edu/nnd.html