BIG DATA CHAPTER 14 SUMMARY

Author
Kevin Loun

Introduction

Chapter 14 discusses various kernel functions used in machine learning for handling objects that cannot be effectively represented by fixed-size feature vectors. This includes complex objects like text documents, molecular structures, and evolutionary trees, where traditional vector representations are inadequate.

Kernel Functions

A kernel function, $\kappa(x, x')$, is defined as a measure of similarity between two objects x and x' in some abstract space X. It is usually required to be symmetric and non-negative. Different types of kernel functions are introduced, including:

- **RBF Kernels:** The Radial Basis Function (RBF) kernel, or Gaussian kernel, is expressed as $\kappa(x, x') = \exp\left(-\frac{||x-x'||^2}{2\sigma^2}\right)$, where σ^2 is the bandwidth. This kernel is sensitive to the Euclidean distance between x and x'.
- **Polynomial Kernels:** These kernels, of the form $\kappa(x, x') = (\gamma x^T x' + r)^d$, where γ, r, and d are parameters, map the inputs into a polynomial feature space.
- **String Kernels:** Useful for comparing sequences or strings by counting common substrings, potentially using complex matching algorithms like suffix trees.
- **Mercer Kernels:** These satisfy the condition that the corresponding kernel matrix (Gram matrix) is positive definite, allowing the application of Mercer's theorem for theoretical guarantees in kernel methods.

Applications and Extensions

Kernels are extensively used in methods that require measuring similarities without explicitly transforming objects into feature vectors. They are crucial for:

- **Support Vector Machines (SVMs):** Where kernels help in defining a high-dimensional hyperplane to separate classes in a more suitable feature space.
- **Kernel PCA:** Kernel PCA uses kernel functions to perform principal component analysis in an implicitly defined feature space, allowing for non-linear dimensionality reduction.
- **Gaussian Processes:** In Gaussian process regression, kernels define the covariance function of the process, crucial for determining the smoothness and other properties of the function being estimated.

Conclusion

The flexibility of kernel methods makes them powerful tools in machine learning, especially in scenarios where the data does not naturally lend itself to traditional vector-based representations. They facilitate working implicitly in high-dimensional spaces without incurring the computational costs associated with such dimensions directly.