Selecting bandwidth for kernel estimator:

see Chap 20 of Wasserman's "All of Statistics, A Concise Course in Statistical Inference" for various risk estimators and bandwidth selectors.

* caveat: it's missing details like data splitting that are part of the equations, but, that he has in his lecture notes:
https://www.stat.cmu.edu/~larry/=sml/densityestimation.pdf
36-708 Statistical Methods for Machine Learning by Larry Wasserstein, CMU

* end of chapter footnote: For large data sets, and (20.25) can be computed quickly using the fast Fourier transform.

see Chap 6 of Bishop's "Pattern Recognition and Machine Language"

see Section 5 of "A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation"
S. J. Sheather,M. C. Jones, 1991, J.R. Statistic Society B, Volume 53, Issue 3, 1991, Pages 683-690

A fast implementation using FFT
https://kdepy.readthedocs.io/en/latest/introduction.html#Selecting-a-suitable-bandwidth

see wikipedia for a single mode data, gaussian estimate:
https://en.m.wikipedia.org/wiki/Kernel_density_estimation#A_rule-of-thumb_bandwidth_estimator