فردین آیار

4) داده ها باید تعداد مناسبی داشته باشند و در فضای حالت مسئله به خوبی توزیع شده باشند.

داده ها نباید مرتب شده باشند. متغیرهای آن ها تا حد ممکن نباید همبستگی داشته باشند. متغیرهای زائد باید از داده ها حذف شده اند. متغیرها باید تا حد ممکن اسکیل یکسانی داشته باشند و بهتر است بین -1 و 1 نرمال شوند. داده های پرت باید با شیوه مناسبی تشخیص داده شوند با روش های مناسب به آنها رسیدگی شود(حذف کل داده، حذف متغیر پرت، جاگذاری متغیر پرت و...)

کل دادگان باید به صورت مناسبی به چند مجموعه برای آموزش، تست و ارزیابی شبکه تقسیم شود.