Username: Rakesh A **Book:** Mastering Machine Learning with R. No part of any chapter or book may be reproduced or transmitted in any form by any means without the prior written permission for reprints and excerpts from the publisher of the book or chapter. Redistribution or other use that violates the fair use privilege under U.S. copyright laws (see 17 USC107) or that otherwise violates these Terms of Service is strictly prohibited. Violators will be prosecuted to the full extent of U.S. Federal and Massachusetts laws.

Algorithm flowchart

The purpose of this section is to create a tool that will help you not just select the possible modeling techniques but also to think deeper about the problem. The residual benefit is that it may help you frame the problem with the project sponsor/team. The techniques in the flowchart are certainly not comprehensive but are exhaustive enough to get you started. It also includes techniques not discussed in this book.

The following figure starts the flow of selecting the potential modeling techniques. As you answer the question(s), it will take you to one of the four additional charts:

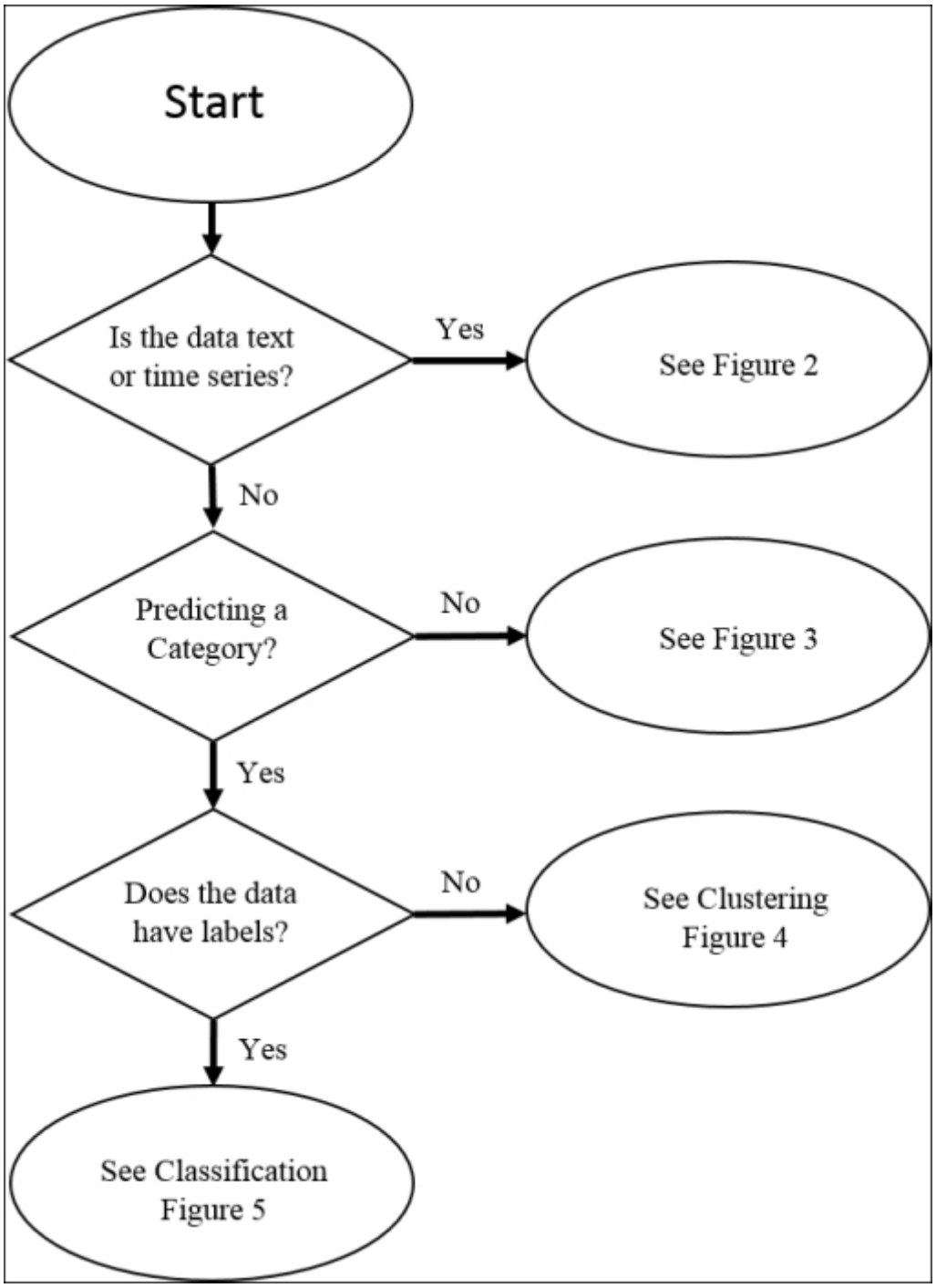

Figure 1

If the data is a text or in the time series format, then you will follow the flow in the following figure:

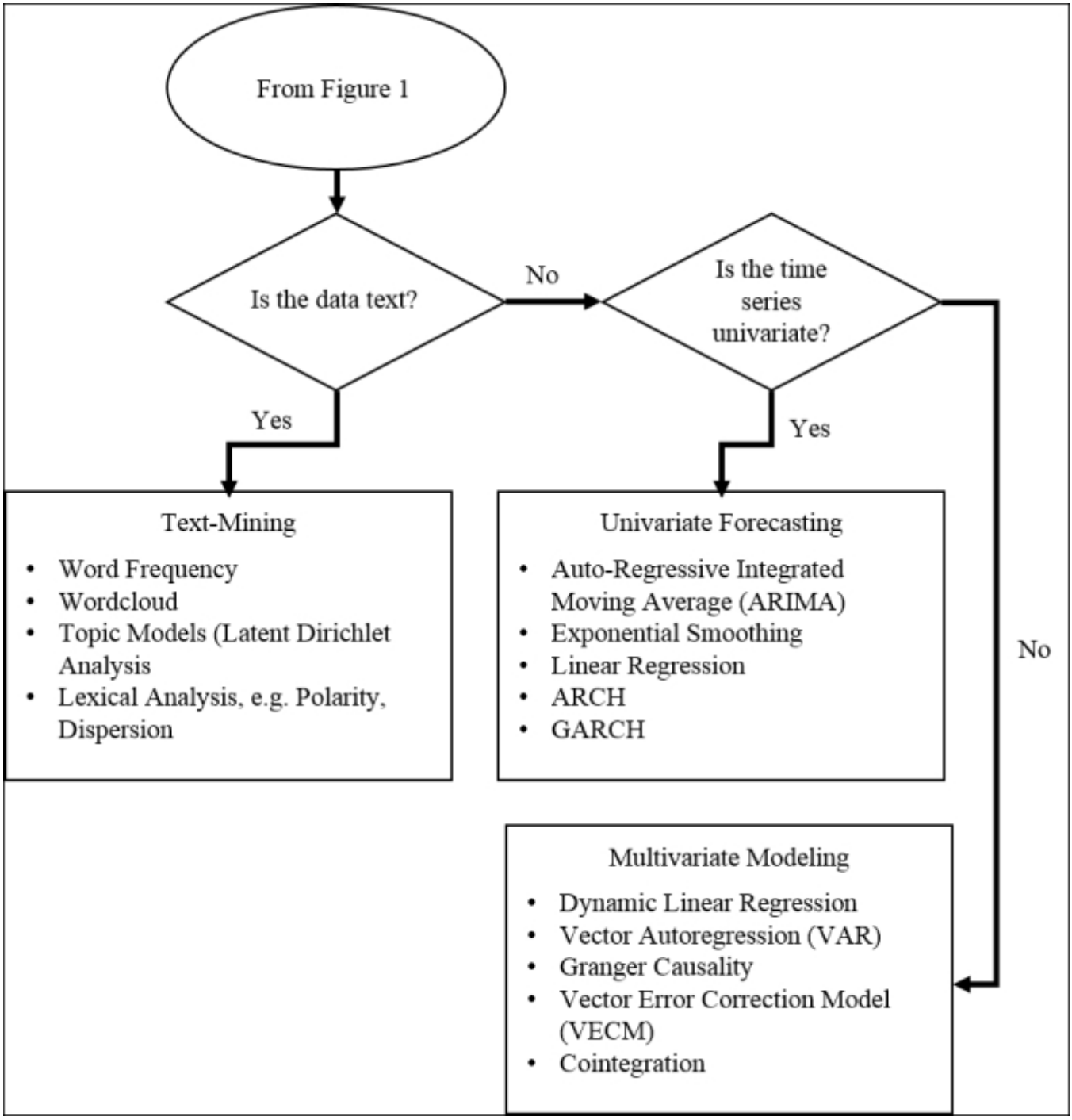

Figure 2

In this branch of the algorithm, you do not have a text or the time series data. Additionally, you are not trying to predict what category the observations belong to.

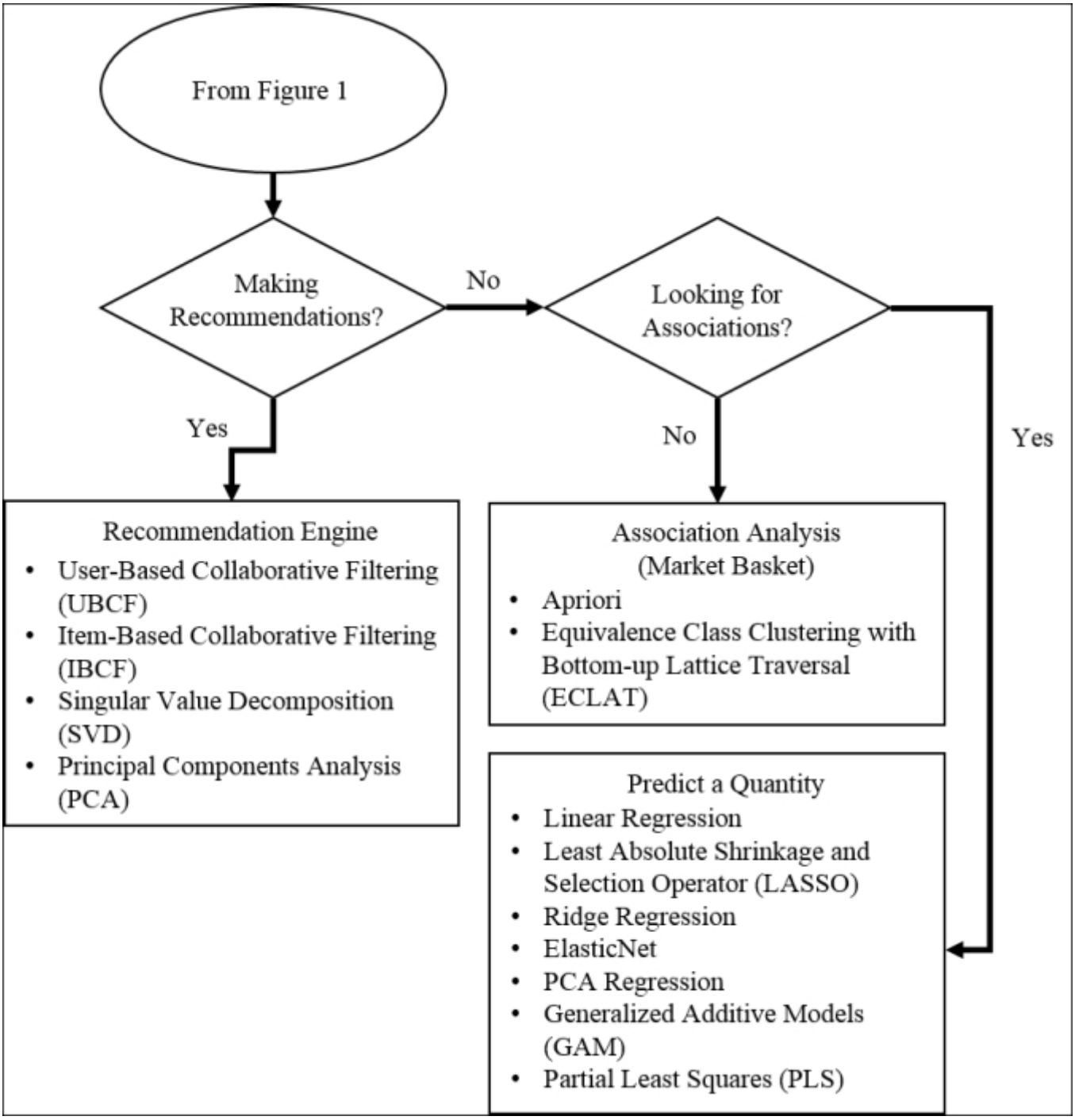

Figure 3

To get to this section, you would have data that is not text or time series. You want to categorize the data, but it does not have an outcome label, which brings us to clustering methods, as follows:

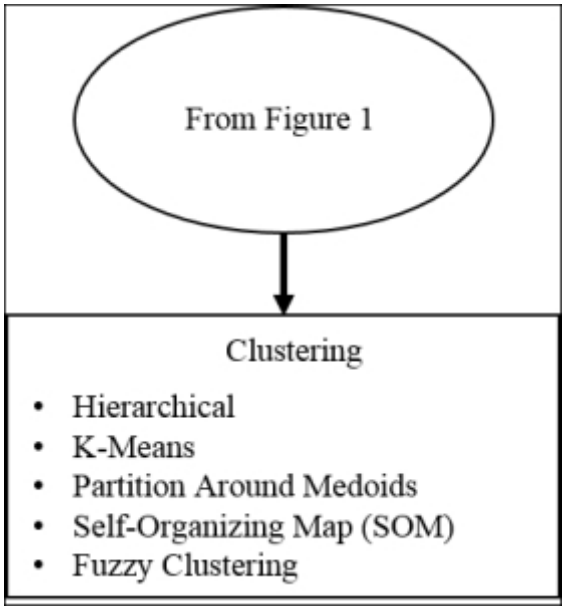

Figure 4

This brings us to a situation where we want to categorize the data and it is labeled, that is, classification:

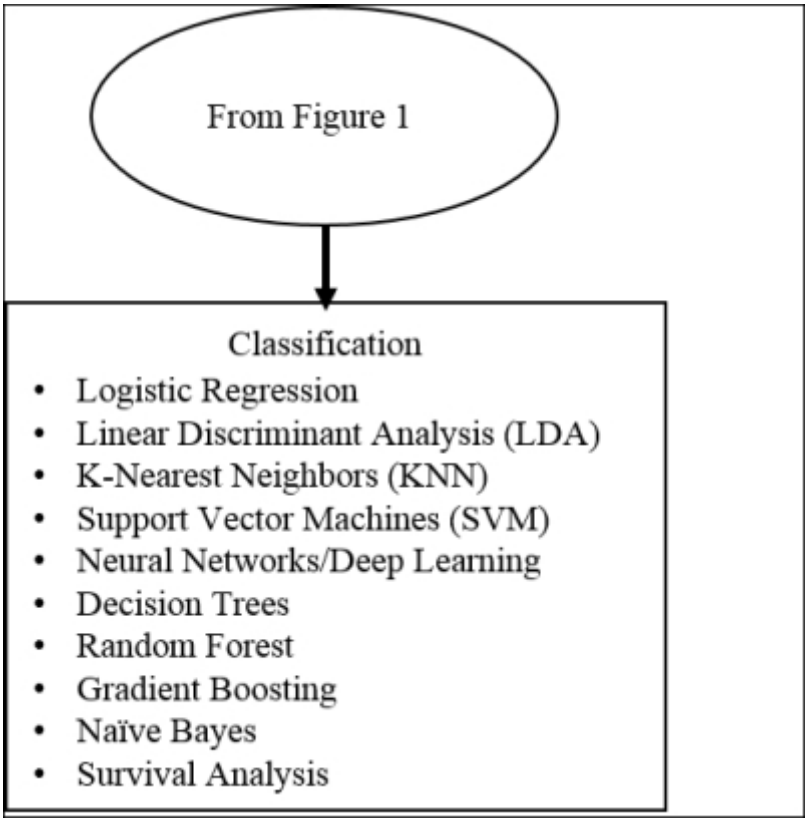

Figure 5