

IAB-DiskAB

Einladung

Imputation von Wirtschaftszweig-Klassifikationen im Betriebs-Historik-Panel auf Basis von *Classification Trees* (CART)

Birgit Pech (ehemals KEM)

CART ist eine Technik der rekursiven Partitionierung, die ursprünglich zu Vorhersagezwecken entwickelt wurde, jedoch auch als Vehikel für die Imputation fehlender Klassifikationsdaten genutzt werden kann. Dies wurde für die Imputation fehlender Werte der Klassifikation der Wirtschaftszweige im Betriebs-Historik-Panel ausgenutzt.

Konkreter wurden solche Entscheidungsbaum-Modelle für die Schätzung der bedingten Auftretswahrscheinlichkeit der Zielkategorien genutzt, wobei nicht nur die jeweilige Ausgangsklassifikation, sondern auch weitere informationstragende Prädiktoren berücksichtigt wurden. Letzteres ist ein wesentlicher Unterschied zu bis dahin am IAB praktizierten Verfahren.

Unter Berücksichtigung der Panel-Struktur wurde ein entsprechendes CART-Verfahren konzipiert, implementiert und evaluiert. Im Zuge der Evaluation wurde die Qualität des neuen Verfahrens mit der Qualität der bisher genutzten Methode verglichen. Die Evaluationsergebnisse lassen darauf schließen, dass die Nutzung von CART-Modellen ein vielversprechender Ansatz für die Imputation fehlender Werte der Wirtschaftszweig-Klassifikationen im Betriebs-Historik-Panel ist.

Montag, 15. Mai, 2017

10:00 Uhr

Raum Re100 E10