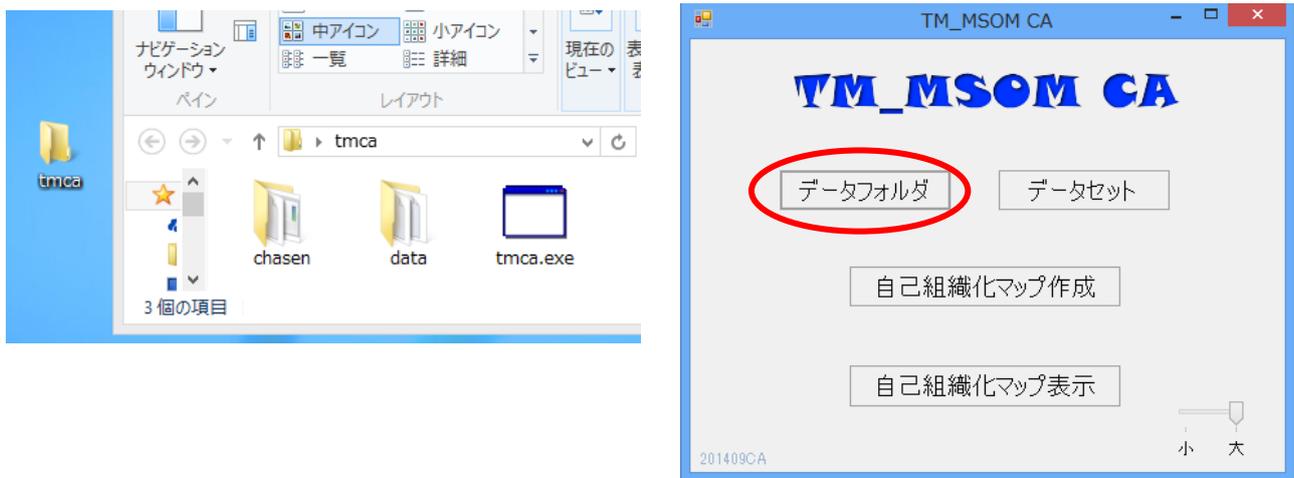


TM_MSOM CA(カテゴリー付き・テキストマイニングの操作解説)

システムをダウンロードし「tmca.zip」ファイルを解凍し、任意の場所に「tmca」フォルダーをおきます。インストラーはありません。不要になった場合は、フォルダーごと削除すれば、システムは削除されます。

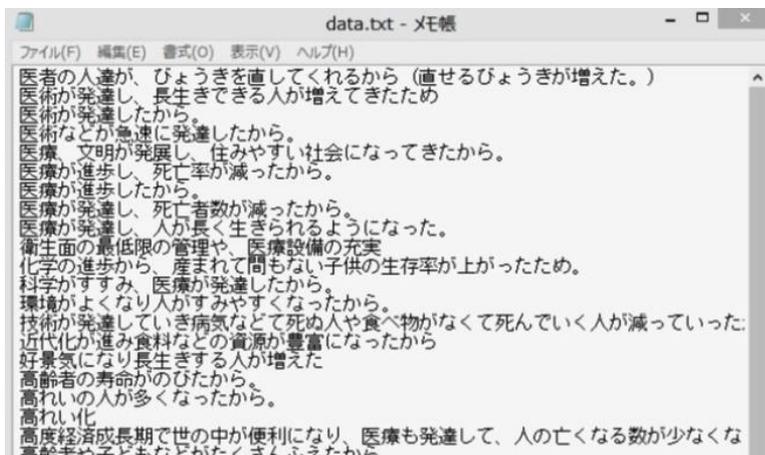
分析したい自由記述のテキストを準備します。一つの文章のかたまりを**改行せずに**テキスト化してください。改行から改行までが一つの単位になります。また、その一つ一つの文章に対応するように、「1～5」の数字のカテゴリー分けを一行ずつ改行したファイルを準備してください。

「tmca」フォルダーの「tmca.exe」ファイルをクリックするとシステムは起動します。システムの「データフォルダ」をクリックして、フォルダーを開きます。

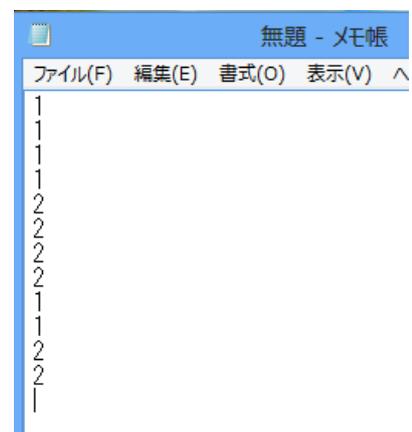


このフォルダーにある「data.txt」ファイルを開き、先ほど準備した分析したい文章を貼り付けます。先に述べたように、カテゴリー分けした一つの文章のかたまりの中では改行はしないようにしてください。約 1000 の文章のかたまりまで分析可能です。次に「idata.txt」のファイルを開き、一つの文章のかたまりに対応するように、準備したカテゴリーの内容を貼り付けて保存してください。カテゴリーは最大5までで、カテゴリー番号は「1～5」の数字で示してください。

「data.txt」のデータセット例



「idata.txt」のデータセット例



データが準備できれば、「データセット」をクリックします。「終了！」が表示されるまで待ちます。次に、右下の「小ー大」のトラックバーで、「小」または「大」を選択します。これは、自己組織化マップのサイズを設定するもので、「小」にすると文章の配置される距離が近くなり重なるケースも多くなります。通常は「大」でよいと思われませんが、集約や重なりを見てくような場合は「小」にするとよいといえます。次に「自己組織化マップ作成」をクリックします。文章量にもよりますが、かなり時間がかかる場合があります。「終了！」が表示されるまで待ち、「終了！」が表示されれば、自己組織化マップが作成されています。作成されたら「自己組織化マップ表示」をクリックすると、自己組織化マップが表示されます。一度作成されれば、アプリケーションを閉じて、新たに起動し「自己組織化マップ表示」をクリックすれば、作成された自己組織化マップが再び表示できます。



自己組織化マップでは、配置された記述内容がカテゴリーに応じて色分けされ配置されます。色のついたセルをクリックするとその記述内容が右に表示されます。同じカテゴリーの文章が複数入っても、そのカテゴリーの色のままです。異なるカテゴリーの文章が一つのセルに入った場合は「■」の色になります。したがって、「■」の色のところは、どのカテゴリーか、セルをクリックして、文章の前



に示された文章番号で判断してください。「TM_MSOM CA」では、文章番号は、「data.txt」のテキストに収められた順番についています。したがって、予めカテゴリごとに、文章を順番に並べておくと、何番から何番まではカテゴリ1というように判断しやすいため、順番に並べておくとよいと思われます。

色のついたセルをクリックすると、記述内容が右枠に表示されます。セルをクリックして、表示された文章を見て、それを抽出したいようであれば、「テキスト保存」のボタンを一度クリックします。これを繰り返します。「抽出テキストの表示」をクリックすると、先ほど抽出した文章がすべて表示されます。この抽出した内容が必要であれば、名前を変更して任意の場所にファイルを保存してください。抽出した内容をすべて削除したい場合は、「抽出テキストの削除」をクリックしてください。