

34) ДОДАТОК А

Набір даних з описами книжкових видань

35) ДОДАТОК Б

Программный код для проведения кластеризации данных

```
%Ввод данных

[num,txt,row] = xlsread('Books_final.xls');
FileNames=txt(:,1);
%Формирование массива данных из массива ячеек

SizeNum2=size(num,2);
FD=zeros((size(num,1)),SizeNum2);
for j=1:(size(num,1));
    for i=1:(size(num,2));
        FD(:,i)=num(:,i);
    end;
end;

%Формирование обучающей и тестовой выборок в соотношении 70%/30% - 140/59

NN=(randperm(199))';
Nt=NN(65:123,1);
Nl=vertcat(NN(1:64,1),NN(124:199,1));
SizeNum1=size(Nl,1);

for i=1:SizeNum1;
    DataLearn(i,:)=FD(Nl(i,1),:);
    FileNamesLearn(i,:)=FileNames(Nl(i,1),:);
end;

%Формирование тестовой выборки 40% от исходного массива
for i=1:size(Nt,1);
    DataTest(i,:)=FD(Nt(i,1),:);
    FileNamesTest(i,:)=FileNames(Nt(i,1),:);
end;

nClusters=9;

%кластеризация k-means
[cluster_idx cluster_center]=kmeans(DataLearn, nClusters, 'distance',
'sqEuclidean', ...
'Replicates',3);

figure(1),
[silh1,h] = silhouette(DataLearn,cluster_idx,'sqEuclidean');
h = gca;
h.Children.EdgeColor = [.8 .8 1];
xlabel 'Silhouette Value', ylabel 'Cluster';

cluster1 = mean(silh1)

E = evalclusters(DataLearn,'kmeans','DaviesBouldin','klist',[1:nClusters])
figure(2),plot(E);

%Проверка на тестовой части выборки

ClustTest=zeros((size(Nt,1)),25);

for id=1:size(Nt,1)
```

```
for ic=1:nClusters
    Ddd(ic)=norm(DataTest(id,:)-cluster_center(ic,:));
end;
[M, INC]=min(Ddd);
NumClust(id,1)=INC;
ClustTest(id,2)=INC;
ClustTest(id,3:25)=cluster_center(INC,:);
end;

ClustTest=num2cell(ClustTest);
ClustTest(:,1)=FileNamesTest(:,1);
```

36) ДОДАТОК В

Значення параметрів опису книжкових видань, які відповідають центрам знайдених дев'яти кластерів

