

**Федеральное государственное бюджетное образовательное  
учреждение высшего образования  
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА  
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ  
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**

**Горшкова Т.Г., Турунцева М.Ю.**

**Теоретические подходы к прогнозированию  
региональных макропоказателей**

**Москва 2020**

**Аннотация.** Работа посвящена анализу существующих теоретических моделей прогнозирования региональных макропоказателей и исследованию возможности прогнозирования российских данных на основе выбранных теоретических подходов. Проведен сравнительный анализ теоретических подходов моделирования региональных данных. Рассмотренные подходы включают в себя индексы диверсификации, основанные на различных экономических теориях, анализ возможности применения композитного индекса благосостояния в качестве прокси переменной для экономического положения и применение к региональным данным динамических и нелинейных моделей. Исследование проводилось на данных по набору макропоказателей (ИПЦ, ВРП на душу населения, уровень безработицы, среднедушевые доходы и пр.) во всех регионах России, а также для регионов, объединенных по федеральным округам и по кластерам, определенным на основе теоретических подходов.

На российских данных были проанализированы различные индексы диверсификации, а также построены ансамбли нейронных сетей и векторные авторегрессии, в том числе учитывающие пространственную зависимость между показателями.

Горшкова Т.Г., научный сотрудник лаборатории макроэкономического прогнозирования ИПЭИ Российской академии народного хозяйства и государственной службы при Президенте РФ.

Турунцева М.Ю., заведующий научно-исследовательской лабораторией макроэкономического прогнозирования ИПЭИ Российской академии народного хозяйства и государственной службы при Президенте РФ.

Данная работа подготовлена на основе материалов научно-исследовательской работы, выполненной в соответствии с Государственным заданием РАНХиГС при Президенте Российской Федерации на 2019 год.

## 1 ВВЕДЕНИЕ

При выборе наиболее эффективной экономической политики государств, состоящих из большого количества субъектов с уникальными географическими и экономическими характеристиками, таких как Российская Федерация, необходимо учитывать особенности всех регионов. Однако в связи с трудоемкостью определения потребностей каждого региона в отдельности и объединении их при выработке общей политики, представляется эффективным рассматривать экономическое положение не каждого региона, а нескольких кластеров, в которые входят регионы с похожими характеристиками.

Актуальность данной работы связана с тем, что вопрос прогнозирования региональных показателей и, в частности, учета пространственной корреляции оказывает значительное влияние на проведение макроэкономической политики на государственном уровне. В научной литературе есть ряд работ, посвященных разработке теоретических подходов к региональному прогнозированию, учитывающих как факторы, единые для любых субъектов исследования, так и уникальные особенности стран и регионов. Подробное изучение имеющихся теоретических моделей поможет выбрать или сформировать подход, наилучшим образом отражающий российскую специфику и учитывающий особенности российской геополитической и экономической ситуации.

В данной работе исследуются возможности применения ряда методов, в том числе методов машинного обучения, для анализа экономической дифференциации между регионами и для прогнозирования региональных показателей, а также применимость данных методов к российским данным.

## 2 ОБЗОР ТЕОРЕТИЧЕСКИХ ПОДХОДОВ К МОДЕЛИРОВАНИЮ РЕГИОНАЛЬНОЙ ДИНАМИКИ И ЭКОНОМИЧЕСКОЙ ДИФФЕРЕНЦИАЦИИ РЕГИОНОВ

Одним из основных теоретических вопросов, предшествующих моделированию и прогнозированию региональных данных, является вопрос о том, как формируется единое экономическое пространство, и как регионы, составляющие это пространство, связаны между собой.

В работе [1] отмечается, что связанность регионов единого экономического пространства является важным фактором, определяющим динамику регионального развития и экономический рост в стране в целом через внутренний экспорт. Кроме того, в данной работе также предполагается, что для успешного развития регионов необходима экономическая политика, направленная на усиление межрегиональной связанности и развитие интеграционных связей. В связи с этим авторы обращают внимание на то, что связи между отдельными показателями в разных регионах могут с разными знаками влиять на показатели социально-экономического развития не только этих регионов, но и соседних, и таким образом, для принятия решений в области региональной экономической политики властям необходимо иметь наиболее подходящие модели и алгоритмы, учитывающие пространственную взаимозависимость.

Существует большое количество работ (см., например, [2], [3], [4], [5]), посвященных определению экономического пространства в целом и региона как субъекта этого пространства в отдельности. В данных работах выделяют наиболее важные, по мнению авторов, свойства региона. В работе [2] к важнейшим структурообразующим свойствам экономического пространства были отнесены плотность, размещение и связанность. Автор [3] в качестве важнейших свойств отмечает организованность, управляемость и конкурентность. В работе [4] дополнительно выделены динамичность, неоднородность (гетерогенность) и асимметричность, а автор [5] отмечает также самоорганизацию.

К настоящему моменту разработано большое число показателей и статистических тестов, позволяющих измерить степень экономической диверсификации регионов и определить, влияет ли данная диверсификация на

экономическую ситуацию и экономическую стабильность (см., например, [6], [7], [8]). Для анализа экономической стабильности часто используются показатели, основанные на региональной безработице и доходах на душу населения.

В работе рассмотрен ряд индексов диверсификации, полученных на основе различных экономических теорий.

Индекс оживления основан на теории организации производства, согласно которой регион, с большим количеством секторов является более разнородным. Для измерения экономической разнородности используются индексы оживления и энтропии.

Индекс оживления рассчитывается по формуле (1):

$$\text{Индекс оживления} = \sum_{i=1}^N \frac{\left(S_i - \frac{1}{N}\right)^2}{1/N}, \quad (1)$$

где  $N$  - количество секторов в экономике,  $S_i$  - доля сектора  $i$  в экономике. В качестве прокси переменной для доли сектора обычно используется доля занятых в данной отрасли.

Чем более одинаково экономическая активность в регионе распределена по секторам, тем больше однородность регионов. При  $N$  секторах одинаковое распределение предполагает, что  $S_i = 1/N$  и индекс оживления равен 0.

Индекс энтропии рассчитывается по формуле (2):

$$\text{Индекс энтропии} = \sum_{i=1}^N S_i \ln\left(\frac{1}{S_i}\right) = - \sum_{i=1}^N S_i \ln(S_i). \quad (2)$$

Индекс энтропии сравнивает существующее распределение показателя (уровень безработицы или доходов в отрасли) по регионам с равно пропорциональным распределением. Чем выше индекс энтропии, тем выше диверсификация, тогда как невысокие значения индекса свидетельствует о специализации региона в определенном виде экономической деятельности.

Индекс Херфинадля обычно используется для определения рыночной концентрации, но может также служить и мерой экономической диверсификации. Данный индекс показывает степень, с которой на экономику определенного региона влияют несколько фирм, и рассчитывается по формуле (3):

$$\text{Индекс Херфиндаля} = \sum_{i=1}^N S_i^2 . \quad (3)$$

Данный индекс принимает значения от 0 до 1, где значение 0 свидетельствует о большом количестве секторов с примерно одинаковой долей в экономике - высокой диверсификации, в то время как значение 1 свидетельствует о слабой диверсификации секторов экономики в регионе.

В работе 1 отмечается, что существует гипотеза о том, что чем больше диверсифицированы сектора в регионе, тем более стабильна экономическая ситуация в данном регионе. Однако авторы утверждают, что результаты проверки этой гипотезы не устойчивы.

Следующие два индекса основаны на экономической теории, в которой основным драйвером регионального экономического роста являются экзогенный спрос и экспорт. Отрасли экономики, влияющие на экзогенный (внешний) спрос называются базовыми отраслями, а отрасли, спрос на которые преимущественно внутренний, считаются небазовыми отраслями. Разница между базовыми и небазовыми отраслями рассчитывается с помощью фактора местоположения по формуле (4):

$$LQ_i = \frac{S_i^{reg}}{S_i^{country}} ,$$

(4)

где  $i$  - сектор экономики,  $S_i^{reg}$  - значение показателя (например, доля занятых) в секторе  $i$  в конкретном регионе,  $S_i^{country}$  - значение показателя в секторе  $i$  в стране в целом. Таким образом, данный показатель сравнивает долю вида экономической деятельности в регионе с соответствующей долей во всей

стране. Если значение индекса больше 1, то в экономике региона соответствующая отрасль имеет большую долю, чем в экономике страны.

Сектора с  $LQ_i > 1$  считаются базовыми (экспортными) секторами и предполагается, что часть продукта, произведенного в таких отраслях, будет экспортирована из региона, в то время как сектора с  $LQ_i < 1$  считаются небазовыми, и их продукты продаются внутри региона.

$LQ$  индекс обычно используется для измерения степени специализации в конкретном секторе и отраслевой концентрации в региональной экономике. Сумма секторальных  $LQ$  - коэффициент специализации - используется как мера региональной специализации. Сумма секторальных  $LQ$ , взвешенных на долю отрасли, называется индексом Хачмана, который рассчитывается по формуле (5):

$$\text{Индекс Хачмана} = \frac{1}{\sum_{i=1}^N \left( \frac{S_i^{reg}}{S_i^{country}} \right) * S_i^{reg}} = \frac{1}{\sum_{i=1}^N LQ_i * S_i^{reg}} . \quad (5)$$

Индекс Хачмана показывает, насколько близки распределения показателя в региональной экономике и в страновой. Данный индекс принимает значения от 0 до 1, где значение 1 свидетельствует о том, что отраслевая структура региона в точности совпадает с отраслевой структурой страны.

Выделяется также ряд индексов, основанных на теории региональных деловых циклов. Согласно данной теории, нестабильность региональной экономики связана с колебаниями спроса на экспорт, особенно на товары роскоши. Предполагается, что экономическая нестабильность может быть объяснена комбинацией стабильных и нестабильных секторов. Для проверки этой гипотезы региональная доля в стабильных и нестабильных секторах используется в качестве показателя экономической диверсификации.

Товары длительного пользования обычно имеют высокую краткосрочную эластичность спроса по доходу и, следовательно, в регионах происходят тем более сильные циклические колебания, чем больше доля товаров длительного пользования в экспорте региона или чем выше доля занятых или дохода в секторах таких товаров и там меньше диверсификация региона.

Предполагается, что чем более региональная отраслевая структура похожа на страновую, тем выше экономическая стабильность. Индекс страновых средних, позволяющий протестировать эту гипотезу, рассчитывается по формуле (6):

$$NAI = \sum_{i=1}^N \frac{s_i^{country} - s_i^{reg}}{s_i^2} \quad (6)$$

Если индекс равен 0, то доля отрасли в регионе и в стране в целом равны. Чем сильнее доля региона отличается от доли национальной экономики, тем выше значение индекса.

Согласно торговой теории, экономические различия возникают вследствие первоначальной наделенности благами, разницы в предпочтениях и конкурентных преимуществах. Торговая теория предполагает, что производственная специализация ведет к экономическому росту. Регионы различаются в результате природных, человеческих и технологических ресурсов, а также инфраструктуры и других пространственных факторов. Институциональные факторы, такие как налоговая структура, окружающая среда и трудовые законы могут влиять на конкурентоспособность региона.

В работе [9] анализируется гипотеза о том, что различия в экономическом росте и благосостоянии регионов связаны с отраслевой специализацией данных регионов. Также в работе отмечается, что прогнозирование региональной динамики усложняется отсутствием единых методик проведения измерений экономических показателей. Авторы видят решение этой проблемы во внедрении единой системы национальных счетов.

Основным фактором, затрудняющим прогнозирование региональных показателей, является, по мнению авторов, значительное число структурных



сдвигов, как на уровне макропоказателей в целом, так и на уровне регионов и отраслей.

Авторы отмечают, что общестрановые прогнозы оказывают значительное влияние на региональные подходы, и на основе данных Министерства экономического развития выделяют набор показателей, которые затем позволят наиболее точно спрогнозировать региональную динамику.

В работе [9] предполагается, что прогноз долгосрочного развития региона учитывает государственный инвестиционный план, и, таким, образом, прогноз общего экономического развития региона может быть разработан в следующей последовательности:

- прогнозируется динамика инвестиций по регионам;
- прогнозируется динамика капиталоемкости производства и объема производства по регионам и отраслям;
- структура промышленного производства;
- динамика потребления домашних хозяйств на основе отношения динамики ВРП рассматриваемого региона к ВВП страны;
- рассчитываются социальные трансферы в регионе, сальдо торгового баланса в регионе и отраслевые дефляторы, основанные на соответствующих общестрановых дефляторах;
- считаются показатели производства, уровень занятости;
- строится прогноз ВРП по ППС.

Таким образом можно построить несколько сценариев экономического развития региона, каждый из которых, по мнению авторов, обеспечивает согласованность и сбалансированность результатов расчетов на общестрановом и региональном уровнях. В результате такого расчета будет получена оценка динамики ВРП.

В статье [9] приведены два возможных сценария регионального развития, зависящие от государственной политики в области распределения инвестиций между субъектами федерации. Согласно первому сценарию распределение инвестиций между округами будет происходить по той же схеме, что и сейчас: пропорции инвестирования не поменяются. Согласно второму сценарию, выделяются так называемые «макрорегионы» - крупные и экономически растущие субъекты, - и пропорции государственных инвестиций будут изменены таким

образом, чтобы уровень экономического развития в субъектах сходил к уровню развития в «макрорегионах».

Для построения региональных сценариев и прогнозов авторы используют общестрановой прогноз МЭР по основным ключевым макропоказателям вплоть до 2030 г.

По результатам проведенных расчетов авторы приходят к выводу, что для второго сценария достижение прогнозных значений показателей по стране в целом, которые являются суммой региональных показателей, возможно только при существенном изменении пропорции инвестиций. В свою очередь при сохранении доли инвестиций неизменной (сценарий 1) разница между показателями в различных субъектах становится очень высокой, что, по мнению авторов, тормозит экономический рост, а также может угрожать политическому единству страны.

### 3 ПРОГНОЗИРОВАНИЕ РЕГИОНАЛЬНЫХ ПОКАЗАТЕЛЕЙ

При прогнозировании региональных показателей используются данные по 83 регионам России за период с 2002-2017 гг. Прогнозирование показателей для республики Крым и города федерального значения Севастополь не проводится в связи с отсутствием данных до 2014 г.

На основе исследований, описанных в предыдущих главах, был отобран ряд методов, которые, как предполагается, дадут точные результаты при прогнозировании российских данных. В связи с большим количеством макроэкономических показателей, экспертным образом были выделены шесть показателей, обладающие наибольшей разделительной способностью между российскими регионами. Данные показатели были отобраны с помощью фрактального анализа. Процесс отбора детально описан в НИР, выполненной в соответствии с государственным заданием РАНХиГС на 2018 год.

В данной работе исследуются региональные данные по 6 показателям. Список показателей приведен в таблице 1.

Таблица 1 – Показатели для анализа неравенства регионов

| Показатель                       | Доступность данных |
|----------------------------------|--------------------|
| ИПЦ                              | 2001-2018          |
| Средние доходы на душу населения | 2001-2018          |
| ВРП на душу населения            | 2001-2017          |

|                                   |           |
|-----------------------------------|-----------|
| Численность рабочей силы          | 2001-2018 |
| Индекс промышленного производства | 2001-2018 |
| Уровень безработицы               | 2001-2018 |

Источник: Росстат

К указанным показателям были применены индексы диверсификации, рассмотренные в первой главе данной работы. В частности, для каждого показателя были рассчитаны индексы, основанные на теории организации производства: индекс оживления и индекс энтропии, – а также индекс LQ, основанный на экономической теории, в которой основным драйвером регионального экономического роста являются экзогенный спрос и экспорт. В данной работе не рассчитывались индексы, основанные на региональных деловых циклах в связи с достаточно коротким промежутком времени, на котором имеются макроэкономические данные (18 лет), а также в связи с изменением региональной структуры РФ, которая не позволяет проанализировать деловые циклы для ряда регионов.

Индексы оживления и энтропии позволяют сравнивать регионы по количеству предприятий, специализирующихся на различных видах экономической деятельности. Если в секторе  $i$  значение индекса  $S_i = 1/N$ , то есть одинаково во всех регионах, то экономическая активность в регионе распределена по секторам равномерно. Данные индексы позволяют сравнить регионы по равномерности распределения и разделить специализированные и неспециализированные регионы. Предполагается, что экономическое положение неспециализированных регионов зависит от общеэкономической ситуации, в результате чего динамика основных макропоказателей в таких регионах будет похожа. В то же время благосостояние специализированных регионов будет сильно зависеть от изменения условий в отрасли их специализации.

Индекс LQ сравнивает отраслевую структуру региона с общестрановой структурой. Если значение индекса больше 1, то в экономике региона соответствующая отрасль имеет большую долю, чем в экономике страны.

Все индексы, приведенные в четвертой главе, рассчитывались по состоянию на 2017 г. На их основе можно определить регионы, значения макроэкономических показателей в которых достаточно похожи друг на друга. На основе индексов

диверсификации можно выделить кластеры регионов, в которых экономическое положение предполагается примерно одинаковым.

Далее более сложные модели применяются к региональным кластерам, выделенным на основе индексов диверсификации, а также определенных с помощью специально разработанной карты кластеризации.

В качестве более сложных моделей, позволяющих спрогнозировать региональные данные, в данной работе используются ансамбль нейронных сетей и пространственные векторные авторегрессии.

Ансамбль нейросетей был построен по аналогии с ансамблем, описанным в работе [10]. Использование нейросетей дает хорошие результаты при моделировании зашумленных данных, а также при использовании данных, количества которых недостаточно для построения более сложных структурных моделей.

Нейросети представляют собой метод, принципиально отличающийся от более известных структурных моделей. Одна из особенностей нейросетей заключается в том, что не важно, какие изначальные данные будут поданы на вход модели. Алгоритм многократно обрабатывает входные параметры и в качестве результата выдает целевой показатель, не имеющий экономического смысла. Таким образом, существенное значение на результаты работы алгоритма имеет спецификация входных данных. Например, если на вход нейронной сети подать одновременно значения ВРП в тыс. рублей и значения ИПЦ в % к предыдущему периоду, вклад ИПЦ в финальном результате будет практически незаметен, так как значение ВРП гораздо больше. В связи с этим при построении персептрона необходимо нормализовать все входные параметры и приводить их к единой размерности. Кроме того, принимая во внимание тот факт, что нейронная сеть не учитывает экономический смысл входящих данных, а также то, что она дает хорошие результаты при работе с большим количеством данных, в данной работе было принято решение строить ансамбль нейросетей не для каждого макроэкономического показателя в отдельности, а для композитного индекса региональных показателей, отражающего общее экономическое положение региона. Данный индекс представляет собой среднее нормализованных значений

рассматриваемых в работе макропоказателей. Таким образом, в качестве входных параметров в нейросеть использовались текущее и запаздывающие значения полученного композитного индекса.

В качестве второго метода получения точных региональных прогнозов использовались векторные авторегрессии. Преимущество VAR заключается в том, что этот тип моделей позволяет учитывать динамику не только одного исследуемого показателя, но и других макропоказателей, которые могут оказывать влияние на исследуемый. Необходимо отметить, что в связи с тем, что по регионам доступны только годовые данные, и, таким образом, модели будут строиться на 16 точках, включение большого числа объясняющих переменных в модель снизить качество полученных оценок. Кроме того, композитный индекс, используемый при построении нейронных сетей, не рекомендуется использовать в качестве объясняемой переменной при построении структурной VAR модели, так как он представляет собой усреднение всех исследуемых в работе макропоказателей. В связи с этим в работе было решено строить систему уравнений для двух показателей: ИПЦ и ВРП, в качестве объясняющих переменных для которых использовались их лаги и остальные макроэкономические показатели в различных комбинациях. Данные модели оцениваются методом обобщенных моментов.

Помимо стандартных VAR в данной работе также строятся пространственные VAR, в которых в качестве объясняющих переменных также используются значения исследуемых показателей в соседних регионах, взвешенные на веса «соседей». При построении этих моделей используется весовая матрица «соседей», в которой значение 1 присваивается тем регионам, которые имеют общую границу, и 0 – тем, которые не имеют.

Таким образом, в данной работе рассматривается три типа VAR:

- стандартные панельные VAR, в которых значение объясняемой переменной в момент времени  $t$  зависит от значений объясняющих переменных в том же регионе в моменты времени  $t - \tau$  ( $\tau \geq 0$ ) ;

- модели, соответствующие структуре стандартных VAR моделей, в которых значение объясняемой переменной в момент времени  $t$  зависит от значений

переменных в соседних регионах (пространственных показателей) в тот же момент времени  $t$ ;

– модели, соответствующие структуре стандартных панелей, в которых значение объясняемой переменной в момент времени  $t$  зависит от значений пространственных показателей в моменты времени  $t - \tau (\tau > 0)$ .

#### 4 ИНДЕКСЫ ДИВЕРСИФИКАЦИИ

В качестве предварительного анализа к основным рядам макропоказателей применены индексы диверсификации, основанные на различных теориях.

Индекс оживления рассчитывается по формуле (7):

$$\text{Индекс оживления} = \sum_{i=1}^N \frac{\left(S_i - \frac{1}{N}\right)^2}{1/N}, \quad (7)$$

где  $N$  – количество секторов в экономике,  $S_i$  – доля сектора  $i$  в экономике.

В соответствии с классификацией Росстата, в России выделяется 19 специализаций, приведенных в таблице 2.

Для каждого региона было рассчитано значение индекса оживления за 2017 г. Максимальное значение индекс принимает в г. Москва – 293.1, минимальное – в Ненецком автономном округе – 42.7, что свидетельствует о высокой степени диверсификации отраслевой специализации в Москве, и о небольшой диверсификации в Ненецком АО. В среднем маленькие значения индекса наблюдаются в регионах Сибири и Дальнего Востока, большие – в крупнейших регионах Центральной России, в т.ч. в Санкт-Петербурге, Воронежской области и др. Полученные результаты соответствуют априорным ожиданиям и отражают историческое разделение России на промышленные зоны.

В таблице 2 также приведены регионы с максимальным значение индекса оживления по областям деятельности.

Таблица 2. Отрасли экономики и регионы с максимальным значением индекса оживления в каждой отрасли

| Отрасль экономики                                                                                            | Регионы с максимальным значением индекса оживления                                                  |
|--------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|
| Сельское, лесное хозяйство, охота, рыболовство и рыбоводство                                                 | республика Дагестан, Карачаево-Черкесская республика, республика Саха (Якутия), Сахалинская область |
| Добыча полезных ископаемых                                                                                   | республика Саха (Якутия), Амурская, Магаданская области, Забайкальский край                         |
| Обрабатывающие производства                                                                                  | Владимирская, Ивановская, Калужская области, республика Ингушетия, Кабардино-Балкарская республика  |
| Обеспечение электрической энергией, газом и паром; кондиционирование воздуха                                 | Чукотский АО                                                                                        |
| Водоснабжение; водоотведение, организация сбора и утилизации отходов, деятельность по ликвидации загрязнений | Волгоградская область                                                                               |
| Строительство                                                                                                | республика Ингушетия                                                                                |
| Торговля оптовая и розничная; ремонт автотранспортных средств и мотоциклов                                   | г. Москва, г. Санкт-Петербург                                                                       |
| Транспортировка и хранение                                                                                   | Тюменская область, Приморский край, Магаданская область                                             |
| Деятельность гостиниц и предприятий общественного питания                                                    | г. Москва                                                                                           |
| Деятельность в области информации и связи                                                                    | г. Москва, г. Санкт-Петербург                                                                       |
| Деятельность финансовая и страховая                                                                          | Липецкая область                                                                                    |
| Деятельность по операциям с недвижимым имуществом                                                            | республика Крым, г. Севастополь, Московская и Ленинградская области                                 |
| Деятельность профессиональная, научная и техническая                                                         | г. Москва, г. Санкт-Петербург, Тюменская область                                                    |
| Деятельность административная и сопутствующие дополнительные услуги                                          | г. Санкт-Петербург                                                                                  |
| Государственное управление и обеспечение военной безопасности; социальное обеспечение                        | Чукотский АО, республика Тыва                                                                       |
| Образование                                                                                                  | республики Чечня, Дагестан, Тыва                                                                    |
| Деятельность в области здравоохранения и социальных услуг                                                    | республики Тыва, Мордовия                                                                           |
| Деятельность в области культуры, спорта, организации досуга и развлечений                                    | республики Тыва, Мордовия                                                                           |
| Предоставление прочих видов услуг                                                                            | г. Москва, г. Санкт-Петербург                                                                       |

Источник: Росстат

На основе таблицы 2 можно четко разделить регионы на три кластера: регионы, специализирующиеся на обрабатывающем производстве, регионы с высоким значения индекса операций по недвижимому имуществу (в основном растущие регионы с высоким спросом на жилье) и прочие. Можно ожидать, что модели, построенные для этих кластеров регионов, будут показывать более точные

результаты, чем модели, построенные на кластерах на основе географического положения.

Помимо индекса оживления был также рассчитан индекс энтропии, также основанный на теории организации производства. Индекс энтропии дает результаты, сопоставимые с индексом оживления. Наиболее диверсифицирована по отраслевой деятельности Москва, наименее - Чукотский и Ненецкий автономные округа.

Далее рассчитан индекс  $LQ_i$ , основанный на предпосылке о том, что основным драйвером регионального экономического роста являются экзогенный спрос и экспорт. Отрасли экономики, влияющие на экзогенный (внешний) спрос называются базовыми отраслями, а отрасли, спрос на которые преимущественно внутренний, считаются небазовыми отраслями. Сектора с  $LQ_i > 1$  считаются базовыми (экспортными) секторами и предполагается, что часть продукта, произведенного в таких отраслях, будет экспортировано из региона, в то время как сектора с  $LQ_i < 1$  считаются небазовыми.

В целом распределение регионов в соответствии с индексом  $LQ_i$  довольно однородно и трудно определить регионы, специализирующиеся на определенных областях. В связи с этим можно кластеризовать регионы на основе транспортируемости произведенных товаров: регионы, производящие транспортируемые товары (продукцию сельского хозяйства, результаты обрабатывающего производства) и не нетранспортируемые товары, такие как водоснабжение, обеспечение электроэнергией, торговля.

Таким образом, на основе индексов диверсификации определено несколько наборов кластеров.

Набор 1 делит регионы на основе индекса оживления: регионы, специализирующиеся на обрабатывающем производстве, регионы с высоким значения индекса операций по недвижимому имуществу и прочие.

Набор 2 делит регионы на основе индекса  $LQ_i$ : регионы, производящие экспортируемые товары, и регионы, специализирующиеся на нетранспортируемых товарах.



Также разработан третий набор кластеров регионов с помощью карты кластеризации, разработанной в рамках исследования анализа динамических свойств российских региональных показателей<sup>1</sup>. Карта позволяет выбирать один из трех методов кластеризации и экспертно задавать количество кластеров, а также исключать Москву. Москва может быть исключена из рассмотрения, так как значения большинства макропоказателей в Москве значительно отличается от значений этих же показателей в других регионах. Таким образом, включение Москвы в модель кластеризации зашумляет данные и усложняет разделение других регионов по кластерам.

Кластеризация производится на основе региональных показателей, представленных в таблице 1. Как и в работе [10], на начальном этапе экспертным образом было выделено 5 кластеров, однако разбиение регионов по кластерам производилось не случайным образом, а на основе алгоритма k-средних.

Как и в работе [10], в третий и четвертый кластеры попало больше всего регионов (48.3% и 36% от общего числа регионов соответственно), что составляет более 80% всех регионов. В связи с этим в данной работе предлагается использовать не пять кластеров, а шесть. В таком случае доли наибольших кластеров составляют 39.3% и 34.8%. Карта России с выделенными кластерами представлена на рисунке 1.

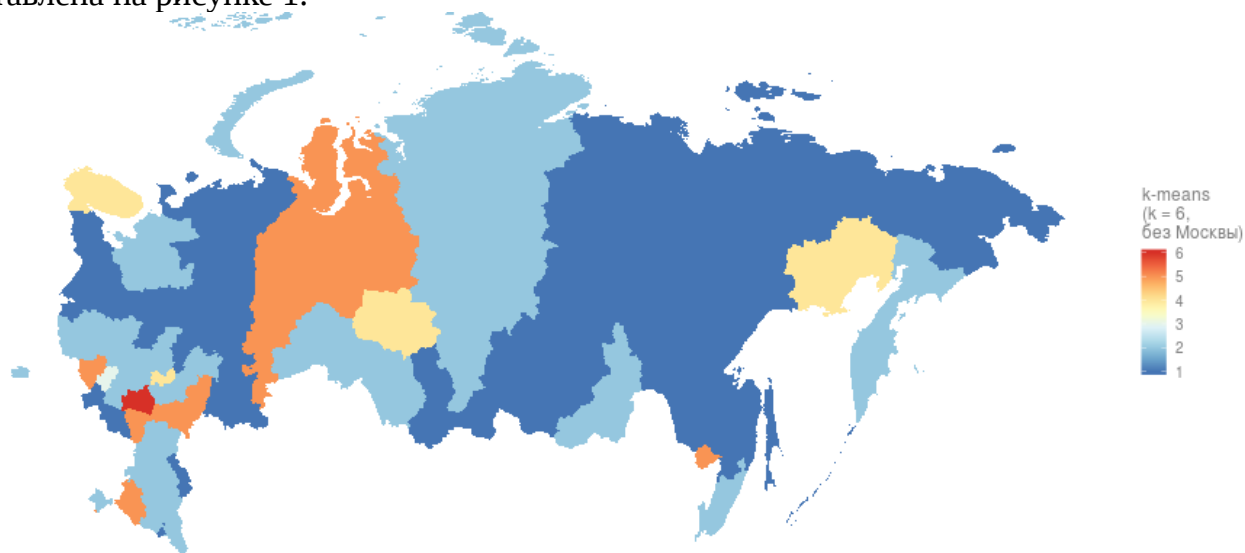


Рисунок 1. Карта России с выделенными кластерами  
Источник: расчеты авторов

<sup>1</sup> [https://lmfiaerranepa.shinyapps.io/russia\\_clustering\\_map/](https://lmfiaerranepa.shinyapps.io/russia_clustering_map/)

## 5 НЕЙРОННЫЕ СЕТИ

Прогнозирование социально-экономического развития регионов на основе шести макропоказателей производится с помощью байесовского ансамбля динамических нейросетей. Так как нейронные сети показывают достаточно точные результаты моделирования на небольших количествах данных, ансамбль строится не на всей доступной статистике, а на данных за 2015-2018 гг. Таким образом, в модель не включается кризис 2014 г, что возможно позволит получить более точные оценки.

На каждый год была построена панель из 89 композитных индексов для каждого региона. Так как используются данные с 2015 год, значения макропоказателей в республике Крым и г. Севастополе также включаются в выборку. Обучающая выборка для построения многослойных персептронов состоит из 160 случайных наблюдений, тестовая выборка – 18 наблюдений. Максимальное число итераций задано на уровне 1000 итераций.

В окончательный байесовский ансамбль включены три персептрона. Параметры нейросетей представлены в таблице 3. Для оценки качества модели рассчитывается внутривыборочная среднеквадратичная ошибка прогноза.

Таблица 3. Параметры ансамбля нейросетей на данных за 2015-2018 гг.

| Нейросеть | Количество скрытых слоев | Количество нейронов в скрытых слоях | Вид активационной функции в скрытых слоях | Вид активационной функции в выходном слое | R MSE    |
|-----------|--------------------------|-------------------------------------|-------------------------------------------|-------------------------------------------|----------|
| 1         | 1                        | 3                                   | гиперболический тангенс                   | линейная                                  | 3<br>2%  |
| 2         | 1                        | 14                                  | гиперболический тангенс                   | линейная                                  | 1<br>2%  |
| 3         | 2                        | 14, 7                               | гиперболический тангенс                   | линейная                                  | 4.<br>8% |

Источник: расчеты авторов

Средняя ошибка ансамбля из трех нейросетей превышает 16%. Однако ошибка третьего персептрона, состоящего из двух скрытых слоев, равна всего 4.8%. На основе полученных результатов можно сделать вывод, что при объединении всех регионов в один кластер и исследовании композитного индекса, составленного из значений макропоказателей, наиболее точные результаты дает многослойный персептрон с двумя скрытыми слоями и 14 нейронами на первом слое и 7 нейронами на втором слое.

Далее ансамбли нейросетей были построены для каждого из шести кластеров из набора 3. Распределение регионов по кластерам описано в таблице 4.

Результаты применения ансамбля нейросетей, описанных в таблице 3, к региональным кластерам приведены в таблице 5. Для кластеров 2 и 6 построение нейросетей не проводилось в связи с малым количеством регионов в кластере.

Таблица 4. Распределение регионов по кластерам

| К<br>ластер | Колич<br>ество<br>регионов<br>/<br>доля кластера | Регионы                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|-------------|--------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1           | 31 / 34.8%                                       | Амурская, Астраханская, республика Башкортостан, Белгородская, Брянская, республика Бурятия, Чукотский АО, Чувашская Республика, Республика Алтай, Иркутская, Ивановская, республика Карелия, Кемеровская, Хабаровский край, республика Хакасия, Кировская, Республика Коми, Костромская, Курская, Ленинградская, республика Марий Эл, Ненецкий АО, республика Северная Осетия - Алания, Оренбургская, Пермский край, Сахалинская, республика Саха (Якутия), республика Тыва, Удмуртская республика, Вологодская, г. Санкт-Петербург                                      |
| 2           | 35 / 39.3%                                       | Алтайский край, Архангельская, Чеченская республика, республика Дагестан, республика Ингушетия, Кабардино-Балкарская республика, Калининградская, Калмыкия, Камчатский край, Карачаево-Черкесская республика, Красноярский край, Курганская, Московская, Нижегородская, Новгородская, Новосибирская, Омская, Орловская, Пензенская, Приморский край, Псковская, Ростовская, Рязанская, Ставропольский край, Свердловская, Республика Татарстан, Тульская, Тверская, Тюменская, Ульяновская, Владимирская, Волгоградская, Ярославская, Забайкальский край, Республика Крым |
| 3           | 2 /<br>2.2%                                      | Калужская, Свердловская                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 4           | 4 /<br>4.5%                                      | Магаданская, республика Мордовия, Мурманская, Томская                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 5           | 12 /<br>13.5%                                    | Челябинская, Ханты-Мансийский АО, Краснодарский край, Самарская, Саратовская, Смоленская, Свердловская, Воронежская, Ямало-Ненецкий АО, Еврейская АО, г. Севастополь, республика Адыгея                                                                                                                                                                                                                                                                                                                                                                                   |
| 6           | 2 /<br>2.2%                                      | Липецкая, Тамбовская                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |

Источник: расчеты авторов

Таблица 5. МАРЕ ансамбля нейросетей для кластеров регионов

| Кла<br>стеры | Нейросеть                 |                            |                                |
|--------------|---------------------------|----------------------------|--------------------------------|
|              | 1 (1<br>слой,<br>нейрона) | 2 (1 слой,<br>14 нейронов) | 3 (2 слоя, 14<br>и 7 нейронов) |
| 1            | 144%                      | 103%                       | 5.2%                           |
| 3            | 0.27%                     | 100%                       | 0.1%                           |
| 4            | 16.8%                     | 52%                        | 63%                            |
| 5            | 143%                      | 63%                        | 69%                            |

Источник: расчеты авторов

Нейронные сети, построенные для кластера 3, дают наиболее точные результаты, однако это может быть связано с тем, что в кластер 3 входит всего 4 региона с очень похожими значениями макропоказателей. Тем не менее, как и для выборки, состоящей из всех регионов России, так и для кластера 1, в который включено наибольшее число регионов, двухслойная нейронная сеть дает достаточно точные результаты – 5.2.%.

На основе двухслойного перспетрона был сделан прогноз композитного индекса экономического состояния регионов на 2019-2020 гг. Далее с помощью самоорганизующейся карты Кохонена, которая представляет собой нейронную сеть, предназначенную для визуализации и кластеризации, на основе полученных прогнозов регионы были разбиты на кластеры. Результаты прогнозируемого разбиения представлены на рисунке 2. При сравнении рисунков 1 и 2 можно заметить, что нейронная сеть прогнозирует незначительное изменение в относительном экономическом положении в Центральной части России - Самарская, Саратовская и соседние с ними области переходят из второго кластера в первый, с более низким значением композитного индекса. Регионы, окружающие Москву, в свою очередь наоборот из первого кластера перемещаются в кластер с более хорошим экономическим положением.

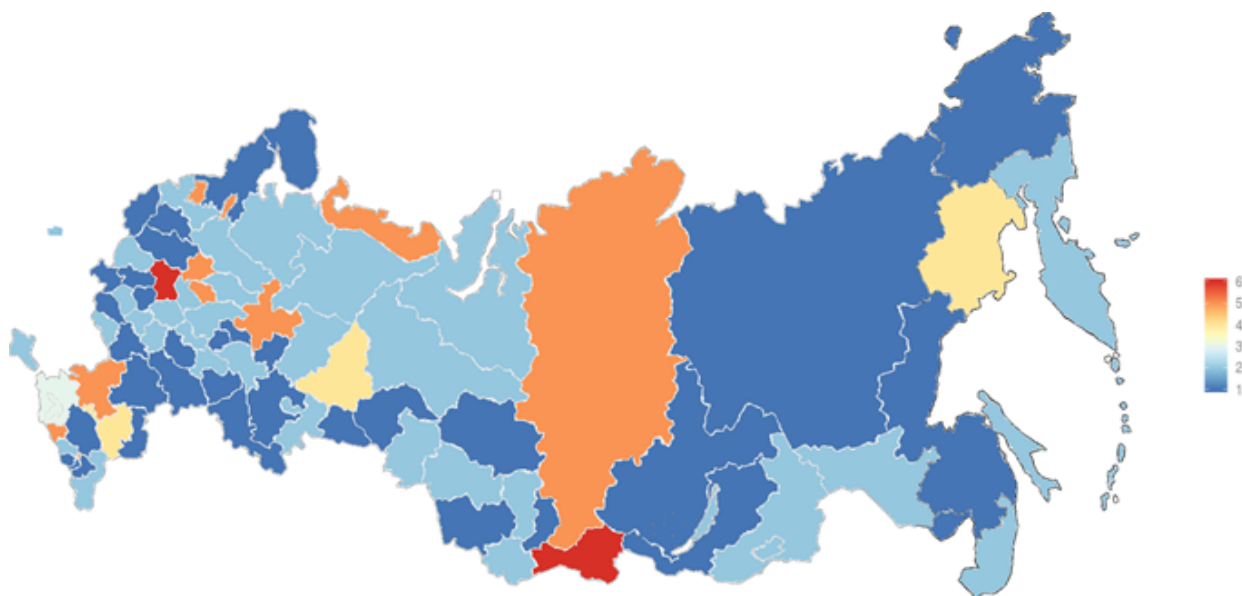


Рисунок 2. Прогнозируемое разбиение на кластеры

Также ансамбли нейросетей были построены для каждого из трех кластеров из набора 2. Распределение регионов по кластерам описано в таблице 6.

Таблица 6. Распределение регионов по кластерам

| Кластер | Количество регионов / доля кластера | Регионы                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|---------|-------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1       | 30 / 38%                            | Белгородская область, Брянская область, Ивановская область, Костромская область, Курская область, Липецкая область, Орловская область, Рязанская область, Смоленская область, Тульская область, Калининградская область, Ленинградская область, Новгородская область, Республика Адыгея, Республика Ингушетия, Кабардино-Балкарская Республика, Республика Северная Осетия - Алания, Ставропольский край, Республика Башкортостан, Республика Марий Эл, Удмуртская Республика, Чувашская Республика, Кировская область, Нижегородская область, Пензенская область, Самарская область, Ульяновская область, Челябинская область, Республика Хакасия, Забайкальский край |

Продолжение таблицы 6

|   |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|---|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2 | 29 / 36.7% | Владимирская область, Калужская область, Московская область, Тамбовская область, Тверская область, Ярославская область, Республика Карелия, Республика Коми, Архангельская область, Вологодская область, Мурманская область, Республика Крым, Астраханская область, Ростовская область, г. Севастополь, Республика Мордовия, Пермский край, Саратовская область, Курганская область, Свердловская область, Республика Бурятия, Красноярский край, Иркутская область, Кемеровская область, Новосибирская область, Омская область, Томская область, Хабаровский край, Еврейская автономная область |
| 3 | 25 / 31.6% | Воронежская область, г. Москва, Ненецкий автономный округ, Псковская область, г. Санкт-Петербург, Республика Калмыкия, Краснодарский край, Волгоградская область, Республика Дагестан, Карачаево-Черкесская Республика, Чеченская Республика, Республика Татарстан, Оренбургская область, Ханты-Мансийский АО, Ямало-Ненецкий АО, Тюменская область, Республика Алтай, Республика Тыва, Республика Саха (Якутия), Камчатский край, Приморский край, Амурская область, Магаданская область, Сахалинская область, Чукотский автономный округ                                                       |

Источник: расчеты авторов

Результаты применения ансамбля нейросетей, описанных в таблице 3, к региональным кластерам приведены в таблице 7.

Таблица 7. МАРЕ ансамбля нейросетей для кластеров регионов

| Кластеры | Нейросеть             |                         |                             |
|----------|-----------------------|-------------------------|-----------------------------|
|          | 1 (1 слой, 3 нейрона) | 2 (1 слой, 14 нейронов) | 3 (2 слоя, 14 и 7 нейронов) |
| 1        | 11.7%                 | 7.3%                    | 17.6%                       |
| 2        | 77.5%                 | 134.3%                  | 46.4%                       |

|   |      |      |      |
|---|------|------|------|
| 3 | 581% | 177% | 398% |
|---|------|------|------|

Источник: расчеты авторов

Нейронные сети, построенные для кластеров 2 и 3, дают очень неточные результаты, для кластера 3 ошибки всех моделей превышают 100, что может быть связано с большой дисперсией индекса экономического состояния в регионах, входящий в данный кластер. Индекс экономического состояния и его приближение, полученное с помощью нейронной сети, приведены на рисунке 3. Помимо неравномерности входящих в кластер регионов, на низкую точность также влияет возможность переобучения нейронной сети, связанная с небольшим количеством регионов в кластере.

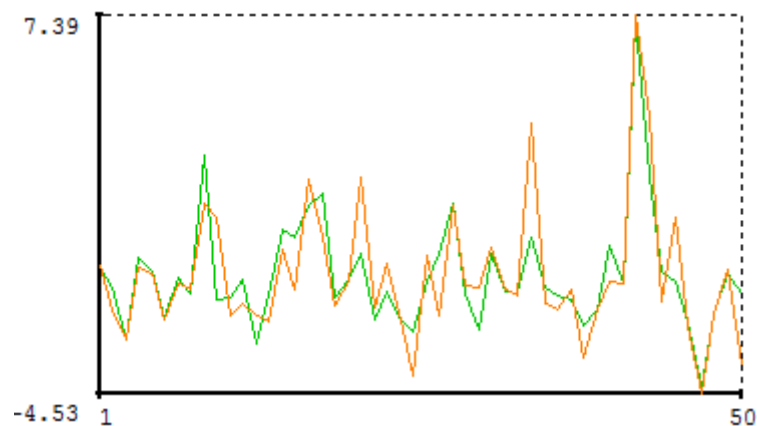


Рисунок 3. Результат нейронной сети для типа нейросети 3, кластера 3.

Оранжевая линия – фактические данные, зеленая линия – результаты нейросети

Также ансамбли нейросетей были построены для двух кластеров из набора 2, рассчитанного на основе индекса  $LQ$ . В первый кластер попали регионы, специализирующиеся на торгуемых товарах, во второй – на неторгуемых товарах. Распределение регионов по кластерам описано в таблице 8.

Результаты применения ансамбля нейросетей, описанных в таблице 3, к региональным кластерам приведены в таблице 9.

Таблица 8. Распределение регионов по кластерам

| Кластер | Количество регионов / доля кластера | Регионы                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|---------|-------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1       | 48 / 59%                            | Белгородская область, Брянская область, Владимирская область, Воронежская область, Калужская область, Костромская область, Курская область, Липецкая область, Орловская область, Рязанская область, Смоленская область, Тамбовская область, Тверская область, Тульская область, Ярославская область, Республика Карелия, Республика Коми, Ненецкий автономный округ, Архангельская область, Вологодская область, Ленинградская область, Мурманская область, Новгородская область, Псковская область, Республика Адыгея, Республика Калмыкия, Республика Крым, Краснодарский край, Астраханская область, Ростовская область, Волгоградская область, Республика Дагестан, Республика Ингушетия, Кабардино-Балкарская Республика, Чеченская Республика, Республика Северная Осетия - Алания, Ставропольский край, Республика Башкортостан, Республика Марий Эл, Республика Мордовия, Чувашская Республика, Кировская область, Оренбургская область, Пензенская область, Саратовская область, Ульяновская область, Курганская область, Республика Алтай, Республика Тыва, Республика Хакасия, Республика Бурятия, Забайкальский край, Красноярский край, Иркутская область, Кемеровская область, Томская область, Республика Саха (Якутия), Камчатский край, Приморский край, Хабаровский край, Амурская область, Магаданская область, Сахалинская область, Еврейская автономная область |

Продолжение таблицы 8

|   |          |                                                                                                                                                                                                                                                                                                                                                                                               |
|---|----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2 | 34 / 41% | Ивановская область, Московская область, г. Москва, Калининградская область, г. Санкт-Петербург, г. Севастополь, Республика Татарстан, Удмуртская Республика, Пермский край, Нижегородская область, Самарская область, Свердловская область, Ханты-Мансийский АО, Ямало-Ненецкий АО, Тюменская область, Чукотский автономный округ, Челябинская область, Новосибирская область, Омская область |
|---|----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Источник: расчеты авторов

Таблица 9. МАРЕ ансамбля нейросетей для кластеров регионов

| Кластеры | Нейросеть             |                         |                             |
|----------|-----------------------|-------------------------|-----------------------------|
|          | 1 (1 слой, 3 нейрона) | 2 (1 слой, 14 нейронов) | 3 (2 слоя, 14 и 7 нейронов) |
| 1        | 6.38%                 | 5.5 %                   | 9.39%                       |
| 2        | 0.6%                  | 1.01%                   | 0.3%                        |

Источник: расчеты авторов

Таким образом, наилучшую точность дают нейронные сети, построенные для кластеров регионов, выделенным на основе индекса  $LQ$ . Все три исследованные нейронные сети дают сопоставимые результаты, однако наибольшая точность

достигается с помощью нейронной сети с двумя скрытыми слоями, что сравнимо с результатами, полученными для кластеров, выделенных на основе карты кластеризации.

Далее по аналогии с моделями для четырех кластеров, на основе двухслойного перспетрона был сделан прогноз композитного индекса экономического состояния регионов для двух кластеров из набора 2 на 2019-2020 гг. Также с помощью самоорганизующейся карты Кохонена на основе полученных прогнозов регионы были разбиты на кластеры. Результаты прогнозируемого разбиения представлены на рисунке 4. При сравнении рисунков 1 и 4 можно заметить, что нейронная сеть прогнозирует незначительное изменение по сравнению с рисунком 1, сделанным на основе кластеров, полученных с помощью карты кластеризации. Прогнозы по второй нейросети незначительно переопределяют регионы в трех центральных кластерах.

При этом список наиболее успешных и отстающих регионов по показателю экономического благосостояния практически не изменился.

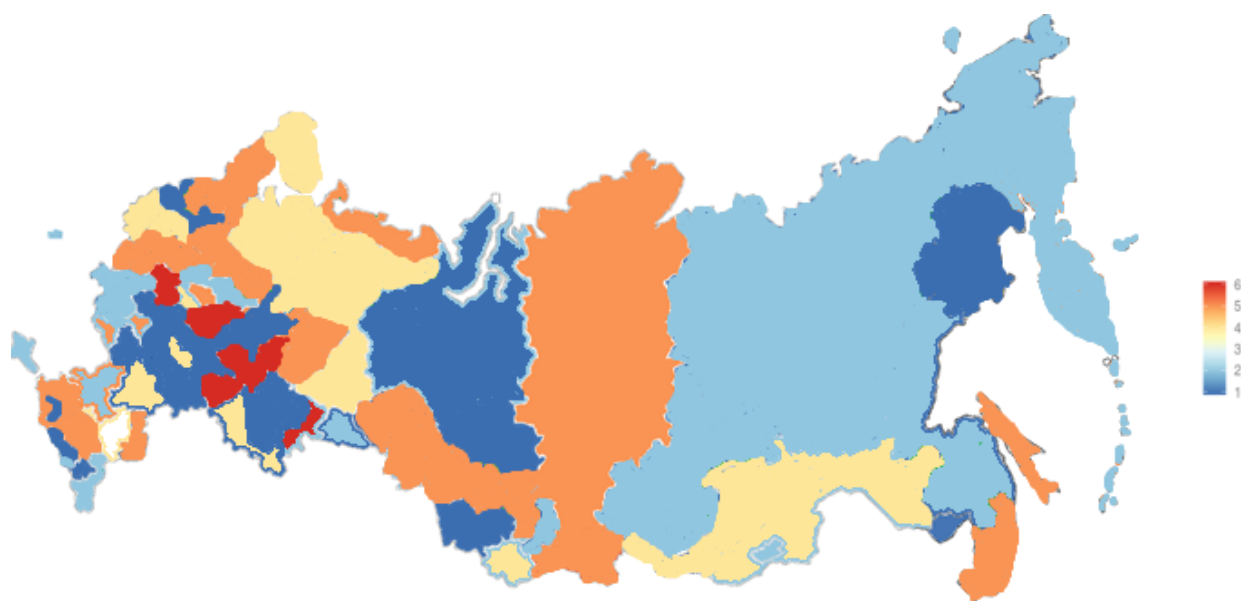


Рисунок 4. Прогнозируемое разбиение на кластеры

Таким образом, по результатам проведенного исследования можно сделать вывод, что для российских данных многослойный персептрон позволяет с точностью около 5% смоделировать экономическое положение регионов. Наиболее точные результаты получаются при использовании всех данных, так как большое количество данных в выборке позволяет нейронной сети выделить зашумленность



в показателях. Также точные результаты получаются для кластеров, состоящих из небольшого количества регионов с очень похожим экономическим положением, однако при маленьком количестве данных существует высокий риск переобучения нейронной сети. В связи с этим в данной работе при построении нейронных сетей предлагается использовать двухслойный персептрон для всей выборки регионов.

## 6 ВЕКТОРНЫЕ АВТОРЕГРЕССИИ

Помимо анализа нейронных сетей, в данной работе по аналогии с работами [11] и [12] строятся векторные авторегрессии. Кроме стандартных VAR, также исследуются модели, учитывающие пространственную зависимость между данными в соседних регионах. Векторные авторегрессии строятся для всех регионов России на данных за 2000-2017 гг. В качестве эндогенных переменных используются значения ВРП и регионального ИПЦ, зависящие от лагов этих же переменных, а также от экзогенной переменной - уровня безработицы.

Стандартная панельная VAR, таким образом, записывается с помощью формулы (8):

$$\begin{cases} \pi_{it} = \alpha_{1i} + \alpha_{2i} Y_{it-1} + \alpha_{3i} \pi_{it-1} + \alpha_{4i} unem_{it-1} + \varepsilon_{it} \\ Y_{it} = \beta_{1i} + \beta_{2i} Y_{it-1} + \beta_{3i} \pi_{it-1} + \beta_{4i} unem_{it-1} + \mu_{it} \end{cases}, \quad (8)$$

где  $\pi_{it}$  - значение инфляции в регионе  $i$  в момент времени  $t$ ,  $Y_{it}$  - значение ВРП в регионе  $i$  в момент времени  $t$ ,  $unem_{it-1}$  - значение безработицы в регионе  $i$  в предыдущий момент времени  $t-1$ .

Модель, соответствующая структуре стандартных VAR моделей, записывается с помощью формулы (9):

$$\begin{cases} \pi_{it} = \alpha_{1i} + \alpha_{2i} Y_{it-1} + \alpha_{3i} \pi_{it-1} + \alpha_{4i} unem_{it-1} + \alpha_5 \sum_{j=1}^N w_{ij} \pi_{jt} + \varepsilon_{it} \\ Y_{it} = \beta_{1i} + \beta_{2i} Y_{it-1} + \beta_{3i} \pi_{it-1} + \beta_{4i} unem_{it-1} + \beta_5 \sum_{j=1}^N w_{ij} Y_{jt} + \mu_{it} \end{cases}, \quad (9)$$

где  $w_{ij}$  - элемент весовой матрицы, характеризующий взаимосвязь между регионами  $i$  и  $j$ .

В работе предполагается, что большее влияние на значение определенного показателя оказывает значение того же показателя, как в исследуемом регионе в предыдущем периоде, так и в соседних регионах. При этом значение второго эндогенного и экзогенного показателей в соседних регионах незначимо влияет на исследуемый показатель. В связи с этим в уравнение для инфляции включены только значения инфляции в соседних регионах и исключены значения ВРП и безработицы.

Модель, соответствующая структуре стандартных панелей, записывается с помощью формулы (10):

$$\begin{cases} \pi_{it} = \alpha_{1i} + \alpha_{2i} Y_{it-1} + \alpha_{3i} \pi_{it-1} + \alpha_{4i} unem_{it-1} + \alpha_5 \sum_{j=1}^N w_{ij} \pi_{jt-1} + \varepsilon_{it} \\ Y_{it} = \beta_{1i} + \beta_{2i} Y_{it-1} + \beta_{3i} \pi_{it-1} + \beta_{4i} unem_{it-1} + \beta_5 \sum_{j=1}^N w_{ij} Y_{jt-1} + \mu_{it} \end{cases} \quad (10)$$

В модели (8) на оба показателя значимо влияют их предыдущие значения, на инфляцию также положительно влияет увеличение уровня безработицы в предыдущем периоде и отрицательно влияет увеличение ВРП в предыдущем периоде, что соответствует экономической теории. На ВРП не влияют ни лаг инфляции, ни лаг безработицы. Согласно тесту Хансена, инструменты, примененные при использовании ОММ, не являются сверхидентифицированными.

При добавлении инфляции и ВРП соседей (модель 9) значение безработицы в предыдущем периоде становится незначимым для инфляции, в то время как лаги инфляции и ВРП в рассматриваемом регионе и текущие значения ВРП и инфляции в соседних регионах значимы на 5% уровне значимости. Для ВРП единственным значимым показателем по-прежнему остается лаг ВРП в рассматриваемом регионе. Согласно тесту Хансена, инструменты, примененные при использовании ОММ, не являются сверхидентифицированными.

При учете в модели не текущих, а запаздывающих значений ВРП и инфляции в соседних регионах (модель 10) результаты моделирования инфляции практически

не изменяются: все показатели, кроме лага безработицы, оказываются значимыми на 5% уровне значимости. Для ВРП, в свою очередь, кроме лага ВРП в рассматриваемом периоде значимыми на 5% уровне значимости становятся лаги ВРП в соседних регионах. Тест Хансена по-прежнему не выявляет сверхидентифицируемость инструментов.

Помимо модели с одной экзогенной переменной – безработицей – также была построена модель, в которой в качестве безработицы используются индекс промышленного производства и среднедушевые доходы населения. Оба показателя также нормализованы.

Однако при построении модели, учитывающей в качестве экзогенного фактора индекс промышленного производства, возникает ошибка, связанная с высоким уровнем сингулярности матрицы. Одной из причин такой ошибки является возможная высокая корреляция между ИПП и ИПЦ, которая приводит к невозможности оценивания параметров модели.

В связи с этим оценивалась модель, в которой в качестве экзогенных переменных использовались запаздывающее значение безработицы и значение среднедушевых доходов. Добавление в модель среднедушевых доходов не улучшает результаты моделирования: для инфляции результаты совпадают с результатами моделей, описанных выше. Для ВРП значимым становится только лаг ВРП в рассматриваемом регионе.

По результатам полученных оценок и, в соответствии с исследованием Baltagi [13], согласно которому даже при отсутствии значимой пространственной связи между регионами, для получения эффективных оценок лучше использовать пространственные модели, для моделирования региональных показателей предлагается использовать модели (8)-(10), в которых учитывается пространственная связь между регионами. Так как из указанных моделей только две модели дают значимые оценки коэффициентов при пространственных показателях, для построения векторной авторегрессии для региональной инфляции и ВРП лучше применять модель, учитывающую в качестве объясняющих переменных первые лаги безработицы, ВРП в соседних регионах и инфляции в соседних регионах. Более детальная спецификация пространственных компонент –

текущие значения или лаги – может быть выбрана с помощью ошибок прогноза модели.

Качество полученных внутривыборочных прогнозов анализировалось с помощью MAPE (средней абсолютной ошибки прогноза, выраженной в процентах), которая рассчитывается по формуле (11):

$$MAPE = \frac{1}{N} \sum_{t=1}^N \frac{|y_t - \hat{y}_t|}{y_t} * 100, \quad (11)$$

где  $y_t$  – фактическое значение прогнозируемого показателя,  $\hat{y}_t$  – спрогнозированное значение.

Несмотря на то, что на основе значимости полученных оценок было выбрано две из четырех моделей для прогнозирования региональных ВРП и ИПЦ, для сопоставимости ошибки прогноза были рассчитаны для всех проанализированных VAR. В таблице 10 приведены значения всех посчитанных ошибок прогноза.

Таблица 10. Ошибки прогнозов VAR моделей

| Модель                                                      | MAPE     |       |
|-------------------------------------------------------------|----------|-------|
|                                                             | инфляции | ВРП   |
| Непространственная (8)                                      | 4,56%    | 2,06% |
| Пространственная с текущими показателями соседей (9)        | 9,17%    | 10%   |
| Пространственная с запаздывающими показателями соседей (10) | 8,67%    | 2,88% |
| Пространственная с учетом доходов                           | 7,87%    | 4,08% |

Источник: расчеты авторов

Из таблицы 10 видно, что модели, учитывающие в качестве объясняющих переменных безработицу и пространственную зависимость между региональными данными, дают наибольшие ошибки прогнозов. При этом для инфляции MAPE ошибки по четырем моделям сопоставимы друг с другом, в то время как для ВРП модель с текущими значениями пространственных показателей в 2-3 раза выше ошибок других моделей. Интересно отметить, что наименьшие ошибки прогнозов

дает непространственная модель, что, однако, не соответствует тому, что оценки коэффициентов при пространственных показателях оказываются значимыми.

По итогам проведенного анализа можно сделать вывод, что для российских данных многослойный персептрон позволяет достаточно точно смоделировать экономическое положение регионов на основе композитного индекса благосостояния. При этом, если в качестве факторов экономического положения, ВВП и инфляция, векторные авторегрессии дают сопоставимые результаты с нейронными сетями.

Среди рассмотренных нейросетей наиболее точные результаты дает двухслойный персептрон, примененный к двум кластерам с большим количеством регионов. Также точные результаты получаются для кластеров, состоящих из небольшого количества регионов с очень похожим экономическим положением, однако при маленьком количестве данных существует высокий риск переобучения нейронной сети. Среди векторных авторегрессий наименьшая ошибка прогноза получается по модели, учитывающей в качестве объясняющих переменных только первые лаги объясняемых переменных и первый лаг безработицы. Однако в моделях, учитывающих пространственную корреляцию между данными, оценки коэффициентов перед пространственными переменными оказываются значимыми. В связи с этим и в соответствии с исследованием Baltagi [13], согласно которому даже при отсутствии значимой пространственной связи между регионами, для получения эффективных оценок лучше использовать пространственные модели, для моделирования региональных показателей предлагается использовать модели, в которых учитывается пространственная связь между регионами. Из трех рассмотренных пространственных моделей наименьшая ошибка прогноза для инфляции получается при добавлении в список объясняющих показателей показателя среднедушевых доходов населения. Наименьшая ошибка прогноза для ВРП получается в модели, где пространственные показатели используются с первым лагом. Так как в модели с учетом дохода оценки коэффициентов при среднедушевых доходах оказываются незначимыми как для инфляции, так и для ВРП, предлагается использовать VAR со следующим набором объясняющих

переменных: первые лаги объясняемых переменных в исследуемом и соседних регионах и первый лаг безработицы.

## 7 ЗАКЛЮЧЕНИЕ

В работе были исследованы методы анализа диверсификации и прогнозирования макроэкономических показателей в регионах России. На первом шаге для ряда макроэкономических показателей были рассчитаны три индекса диверсификации, основанные на различных экономических теориях. По результатам этого анализа было выделено три набора кластеров, к которым применялись различные спецификации нейросетей.

Наименьшие ошибки внутривыборочных прогнозов были получены для кластеров, состоящих из достаточно большого количества регионов. Данный результат может быть связан с тем, что нейронные сети более эффективны для рядов с большим количеством данных. Также точные результаты дают перцептроны, примененные к кластерам, в которые входят регионы с похожим экономическим положением, но в данном случае существует риск переобучить нейронную сеть на слишком однородных данных. Использование большого количества слоев в перцептроне также приводит к переобучению сети. Таким образом, по результатам проведенного исследования можно сделать вывод, что наиболее оптимальной нейросетью является двухслойный перцептрон, примененный к двум кластерам с большим количеством регионов.

Затем на российских данных были построены векторные авторегрессии. При этом композитный индекс, используемый при построении нейронных сетей, не рекомендуется использовать в качестве объясняемой переменной для структурной VAR модели, так как он представляет собой усреднение всех потенциальных объясняющих макропеременных. В связи с этим в работе были построены VAR для двух показателей: ИПЦ и ВРП.

Векторные авторегрессии также дают достаточно маленькие ошибки внутривыборочных прогнозов. В работе исследовалось четыре VAR для ВРП и региональной инфляции:

– без пространственных связей между регионами;

- пространственная, в которой в качестве объясняющих переменных используются текущие значения региональной инфляции и ВРП в соседних регионах;

- пространственная, в которой в качестве объясняющих переменных используются предыдущие значения региональной инфляции и ВРП в соседних регионах;

- пространственная, в которой в список объясняющих переменных также добавлены данные по среднему душевому доходу в рассматриваемом регионе.

Непространственная модель дает более точные значения внутривыборочных прогнозов. Однако в пространственных моделях коэффициенты при показателях соседей оказываются значимыми, в связи с чем для получения несмещенных оценок необходимо учитывать пространственные связи между регионами. Также необходимо отметить, что в модели с учетом доходов оценки коэффициентов при среднедушевых доходах оказываются незначимыми как для инфляции, так и для ВРП, что позволяет не учитывать этот показатель в модели. Из оставшихся моделей наименьшие ошибки прогноза как для ВРП, так и для региональной инфляции, получаются в модели, где пространственные показатели используются с первым лагом.

Таким образом, по результатам проведенного исследования можно сделать вывод, что индексы диверсификации, основанные на экономических теориях, позволяют хорошо разделить регионы по их экономическому благосостоянию, определяемому на основе композитного индекса благосостояния. В свою очередь динамические модели: нейронные сети и векторные авторегрессии – позволяют достаточно точно смоделировать как региональную динамику отдельных макропоказателей, так и общую ситуацию в регионах России.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 Полякова А.Г., Симарова И.С. Концептуальная модель управления развитием региона с учетом уровня пространственной связанности//Экономика региона – 2014 – № 2 (38). С. 32-42.
- 2 Гранберг А.Г., Основы региональной экономики//ГУ–ВШЭ – 2001
- 3 Бияков О.А. Экономическое пространство. Сущность, функции, свойства//Вестник Кузбасского государственного технического университета – 2004 – Вып. 2. Стр. 101-108
- 4 Митрофанова И.В. К вопросу о сущности экономического пространства//Региональная экономика. Теория и практика. – 2008 – No 8. – С. 2-10.
- 5 Моргоев Б.Т. Асимметричная трансформация российского экономического пространства// Ростов н/Д.: Изд-во СКНЦ ВШ – 2005. – 244 с.
- 6 Dr. Khem Raj Sharma. Analysis Division Department of Business, Economic Development and Tourism STATE OF HAWAII//State's Innovation Initiative – 2008.
- 7 Izraeli O., Murphy K.J. The effect of industrial diversity on state unemployment rate and per capita income//The Annals of Regional Science – 2003. – Vol. 37. P. 1-14.
- 8 Siegel P.B., Alwang J., Johnson T.G. Toward an improved portfolio variance measure of regional economic stability//Review of Regional Studies – 1994. – Vol. 24. P. 71-86.
- 9 Gagarina G. Y, Dzyuba E. I., Gubarev R. V, Fayzullin F. S. Forecasting of Socio-Economic Development of the Russian Regions//Economy of region – 2017. – Vol 4. P. 1080-1094
- 10 Узяков М.Н., Сапова Н.Н., Херсонский А.А. Инструментарий макроструктурного регионального прогнозирования: методические подходы и результаты расчетов // Проблемы прогнозирования. 2010. № 2. С. 3-20.
- 11 Beenstock M., Felsenstein D. Spatial Vector Autoregressions//Spatial Economic Analysis/ - 2007. – Vol. 2(2). P. 167-96
- 12 Patuelli R., Mayor M. Short-Run Regional Forecasts: Spatial Models Through Varying Cross-Sectional and Temporal Dimensions//from book From Administrative Spatial Units to Local Labour Market Areas. Some Remarks on the Unit of Investigation of Regional Economics with Particular Reference to the Applied Research in Italy. – 2012. – P.173-192
- 13 Baltagi, B.H., Li, D., 2004. Prediction in the panel data model with spatial correlation. In: Anselin, L., Florax, R.J.G.M., Rey, S.J. (Eds.), Advances in Spatial Econometrics: Methodology, Tools and Applications. Springer, Berlin, pp. 283–295 (Chapter 13).